



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113398

TO: Michael Borin
Location: rem/2a55/2c70
Art Unit: 1631
Wednesday, February 04, 2004

Case Serial Number: 09/582486

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Borin,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:58 ; Search time 21 seconds

(without alignments)

1424.212 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTTPTFSLAELOOGLHOD.....ATFQDWIGNYNIRRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96158682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	1629	99.5	311	2 T52312	deacetoxycephalosporin
2	1627	99.4	311	2 A32043	deacetoxycephalosporin
3	1185.5	72.5	314	2 S30900	deacetoxycephalosporin
4	965.5	59.0	310	2 S40253	deacetoxycephalosporin
5	935.5	57.4	318	2 A39204	deacetoxycephalosporin
6	937	57.2	319	2 S54100	deacetoxycephalosporin
7	926.5	56.6	332	2 A29711	deacetoxycephalosporin
8	852	52.0	313	2 S54101	deacetoxycephalosporin
9	176	10.8	329	2 A61155	isopenicillin N sy
10	164	10.0	329	2 A29894	isopenicillin N sy
11	162	9.9	326	2 S08218	isopenicillin N sy
12	162	9.9	326	2 S54099	isopenicillin N sy
13	146.5	8.9	362	2 D84713	probable dioxigena
14	144.5	8.8	376	2 D96635	probable gibberell
15	143	8.7	333	2 B32344	isopenicillin N sy
16	142.5	8.7	358	2 C84713	probable dioxigena
17	140	8.6	348	2 S33510	flavonol synthase
18	138	8.4	328	2 S15284	isopenicillin N sy
19	137	8.4	321	2 A58458	isopenicillin N sy
20	136.5	8.3	380	2 T48491	gibberellin 20-oxi
21	134.5	8.2	349	2 T07373	flavonol synthase
22	133.5	8.2	331	2 A26467	isopenicillin N sy
23	131.5	8.0	355	2 AG0160	probable iron/asco
24	129.5	7.9	331	2 S04441	isopenicillin N sy
25	125	7.6	382	2 T11847	gibberellin 20-oxi
26	124.5	7.6	370	2 T11849	gibberellin 20-oxi
27	123	7.5	339	2 T05903	iron deficiency pr
28	122.5	7.5	331	2 A27355	isopenicillin N sy
29	122.5	7.5	386	2 T09664	gibberellin 20-oxi

30	122.5	7.5	386	2 T09675	probable gibberell
31	121	7.4	356	2 T05119	leucoanthocyanidin
32	121	7.4	371	2 T08008	leucoanthocyanidin
33	118.5	7.2	377	2 T10222	gibberellin 20-oxi
34	118	7.2	365	2 S31921	naringenin 3-dioxy
35	117.5	7.2	358	2 S44261	SRG1 protein - Ara
36	117.5	7.2	370	2 T47932	1-aminocyclopropan
37	117	7.1	320	2 C83628	probable oxidoredu
38	116.5	7.1	356	2 T05551	SRG1 protein-relat
39	115	7.0	380	2 T06787	gibberellin 20-oxi
40	114.5	7.0	321	2 T41002	hypothetical prote
41	114	7.0	356	2 T07972	leucoanthocyanidin
42	113.5	6.9	338	2 S47972	dioxygenase-iron
43	113	6.9	298	2 T09145	ethylene-forming e
44	113	6.9	352	2 T01606	probable flavonol
45	113	6.9	380	2 T06439	gibberellin 20-dio

ALIGNMENTS

RESULT 1

T52312

deacetoxycephalosporin C synthetase [imported] - Streptomyces clavuligerus

C:Species: Streptomyces clavuligerus

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 24-Aug-2001

C:Accession: T52312

R:Kovacevic, S.; Tobin, M.B.; Miller, J.R.

J. Bacteriol. 172, 3952-3958, 1990

A:Title: The beta-lactam biosynthetic genes for isopenicillin N epimerase and deacetoxyce

A:Reference number: Z26033; MUID:90299822; PMID:1694525

A:Accession: T52312

A:Status: Preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-311 <KO>

A:Cross-references: EMBL:M32324; PIDN:AAA26715.1

C:Genetics:

A:Gene: cefE

C:Superfamily: isopenicillin N synthase

Query Match 99.5%; Score 1629; DB 2; Length 311;
Best Local Similarity 99.7%; Pred. No. 4.9e-140;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MDTTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSE	60
Db	1	MDTTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDLVIDFFEHGSE	60
Qy	61	AEKRAVTSVPVTTMRGFTGLESESTAQTNTGVSYSYSCYSMTADNLPSPGDFGRIT	120
Db	61	AEKRAVTSVPVTTMRGFTGLESESTAQTNTGVSYSYSCYSMTADNLPSPGDFGRIT	120
Qy	121	QYFDQVTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM	180
Db	121	QYFDQVTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM	180
Qy	181	APHYDLSMTLIQOFTPCANGFVSLQAEVGAFTDLPIYRDAVLVFCGATATLVGQVKA	240
Db	181	APHYDLSMTLIQOFTPCANGFVSLQAEVGAFTDLPIYRDAVLVFCGATATLVGQVKA	240
Qy	241	PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG	300
Db	241	PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG	300
Qy	301	NYVNIRRTSKA 311	
Db	301	NYVNIRRTSKA 311	

RESULT 2

A32043

deacetoxycephalosporin C synthetase - Streptomyces clavuligerus

C:Species: Streptomyces clavuligerus

THIS PAGE BLANK (USPTO)

C;Date: 13-Jul-1989 #sequence_revision 13-Jul-1989 #text_change 19-May-2000
C;Accession: A32043
R;Kovacevic, S.; Weigel, B.J.; Tobin, M.B.; Ingolia, T.D.; Miller, J.R.
J. Bacteriol. 171, 754-760, 1989
A;Title: Cloning, characterization, and expression in *Escherichia coli* of the Streptomycin
A;Reference number: A32043; MUID:89123150; PMID:2644235
A;Accession: A32043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <MOV>
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 99.4%; Score 1627; DB 2; Length 311;
Best Local Similarity 99.4%; Pred. No. 7.5e-140;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGE 60
DB 1 MDVTPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGE 60
QY 61 AEKRAVTSPTMRRGFTGLSESTAQITNTGSDYSCYMGTAADNLPSPGDFGRITW 120
DB 61 AEKRAVTSPTMRRGFTGLSESTAQITNTGSDYSCYMGTAADNLPSPGDFGRITW 120
QY 121 QYFPROQTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 121 QYFPROQTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPPOVPEHRSABEQPLRM 180
QY 181 APHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSRTSSVFFLRPNADTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSRTSSVFFLRPNADTFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRTSKA 311
DB 301 NYVNIIRTSKA 311

RESULT 3
S30900
deactoxyccephalosporin C synthetase - Streptomycetes lactamdurans
C;Species: Streptomycetes lactamdurans
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-2000
C;Accession: S30900; S28383
R;Coque, J.J.R.; Martin, J.F.; Liras, P.
Mol. Gen. Genet. 236, 453-458, 1993
A;Title: Characterization and expression in Streptomycetes lividans of cefD and cefE genes
clavuligerus.
A;Reference number: S30900; MUID:93173127; PMID:8437592
A;Accession: S30900
A;Molecule type: DNA
A;Residues: 1-314 <COQ>
A;Cross-references: EMBL:213974; NID:949299; PIDN:CAA78376.1; PID:949300
A;Note: the source is designated as Nocardia lactamdurans
C;Genetics:
A;Gene: cefE
C;Superfamily: isopenicillin N synthase

Query Match 72.5%; Score 1186.5; DB 2; Length 314;
Best Local Similarity 69.8%; Pred. No. 7e-100;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGE 60
DB 3 DATVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGE 62
QY 61 AEKRAVTSPTMRRGFTGLSESTAQITNTGSDYSCYMGTAADNLPSPGDFGRITW 120
DB 63 AEKRAVTSPTMRRGFTGLSESTAQITNTGSDYSCYMGTAADNLPSPGDFGRITW 122

QY 121 QYFPROQTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 123 DYFARMYRASQDVARQVLTSSVGAEPVGMDFDCEPLLRFRYPPEVPEHRSABEQPLRM 182
QY 181 APHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 240
DB 183 APHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 242
QY 241 PRHVAAPRRDQIAGSRTSSVFFLRPNADTFVPLARECGFDVSLDGETATFQDWIGG 300
DB 243 PGHVAAFCADKRVGSSRTSSVFFLRPNADTFVPLARECGFDVSLDGETATFQDWIGG 302
QY 301 NYVNIIRTSKA 311
DB 303 NYVNIIRTSKA 313

RESULT 4

S40253

deactoxyccephalosporin C synthetase - Streptomycetes lactamdurans

C;Species: Streptomycetes lactamdurans

C;Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 19-May-2000

C;Accession: S40253

R;Coque, J.

submitted to the EMBL Data Library, February 1993

A;Reference number: S40253

A;Accession: S40253

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-310 <COQ>

A;Cross-references: EMBL:221687; NID:9438193; PIDN:CAA79803.1; PID:9438194

A;Note: the source is designated as Nocardia lactamdurans

C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 59.0%; Score 965.5; DB 2; Length 310;
Best Local Similarity 61.2%; Pred. No. 7.8e-80;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGE 61
DB 3 DKTVPVFSMAELDGRQDEFRWAR-RGVFLYTGATERDHRVATDTAMDFFAQGTAE 61
QY 62 EKRAVTSPTMRRGFTGLSESTAQITNTGSDYSCYMGTAADNLPSPGDFGRITW 121
DB 62 EKQAVTTKVTMRRGYSALAEASTAQVNTGTVDYMSYMGIGGLNLPSPKESFVWTD 121
QY 122 YFPROQTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPPOVPEHRSABEQPLRM 181
DB 122 YFDSLRYAAQETARLVLTAAQVYDGDLDLDCDPLRLRYFPPEHRAAEYEPERMA 181
QY 182 PHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 241
DB 182 PHYDLSNVTLIQTTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVGGQVKA 241
QY 242 RHVAAAPRRDQIAGSRTSSVFFLRPNADTFVPLARECGFDVSLDGETATFQDWIGG 301
DB 242 NHHVSPDAMKSGDSRTSSVFFLRPNADTFVSPDARKYGLDVSMDMEKATFGDWIGG 301

QY 302 YVNI 305
DB 302 YVTM 305

RESULT 5

A39204

deactoxyccephalosporin C synthetase hydroxylase - Streptomycetes clavuligerus

C;Species: Streptomycetes clavuligerus

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-May-2000

C;Accession: A39204; A23713

R;Kovacevic, S.; Miller, J.R.

J. Bacteriol. 173, 398-400, 1991

A;Title: Cloning and sequencing of the beta-lactam hydroxylase gene (cefH) from Streptomycetes.

THIS PAGE BLANK (USPTO)

Reference number: A39204; MUID:91100311; PMID:1987130

Accession: A39204

Molecule type: DNA

Residues: 1-318 <KOV>

Cross-references: GB:M3809; GB:M37186; NID:g153206; PIDN:AAA26716.1; PID:g153207

Baker, B.J.; Dotzlat, J.E.; Yeh, W.K.

Biol. Chem. 266, 5087-5093, 1991

Title: Deacetoxycephalosporin C hydroxylase of Streptomyces clavuligerus. Purification

Reference number: A23713; MUID:91161600; PMID:2002049

Accession: A23713

Molecule type: protein

Residues: 2-29,92-100 <BAK>

Genetics:

Gene: cefF

Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 57.4%; Score 939.5; DB 2; Length 318;
Best Local Similarity 58.6%; Pred. No. 1.8e-77;
Matches 180; Conservative 43; Mismatches 81; Indels 3; Gaps 1;

2 DTTVPFSLAELOOGLHODEFRRLRDKGLFYLTDCGLTDLTKSAKDLVIDFFEHGSEA 61

3 DTPVPFNLAAUREGADQKRECVTGMGVFYLTVGAGDKDRLATDTAMDFFANGTEA 62

62 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121

63 EKAAVTDVPMRGVSALAEASTAQVTRTGYTSDYSGMISGNVFPSPPEFERVWTE 122

122 YFDRQYTSARAVAREVLRTG---TEPDGGVAFDCEPLLRFPRFPQVPEHRSAREQPL 178

123 YFDKLYAAQETARLVLTASGGYDAEIVGSLDELDDADPVLRLRYPEVPEHRSABEPR 182

179 RNAPHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQV 238

183 RNAPHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQV 242

239 KAPRHVAAAPRRDQIAGSSRTSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWI 298

243 PAPRHVRSFGAGMRGSDRTSSVFLRPTDFSPVAKARSYGLAVLDMETATFGDWI 302

299 GGNVNI 305

303 GTNYVTM 309

RESULT 6

leacetoxycephalosporin C synthetase 1 - Lysobacter lactamgenus (strain YK90)

Species: Lysobacter lactamgenus

Variety: strain YK90

Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-May-2000

Accession: S54100

Kimura, H.; Izawa, M.; Miyashita, H.; Shimizu, Y.; Sumino, Y.; Suzuki, M.

submitted to the EMBL Data Library, October 1990

Description: Gene cluster involved in the cephalosporin biosynthesis from Lysobacter

Reference number: S54099

Accession: S54100

Molecule type: DNA

Residues: 1-319 <KIM>

Cross-references: EMBL:X56660; NID:g769806; PIDN:CAA39984.1; PID:g769808

Experimental source: strain YK90

Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Keywords: antibiotic biosynthesis

Query Match 57.2%; Score 937; DB 2; Length 319;

Best Local Similarity 58.0%; Pred. No. 3.1e-77;

Matches 177; Conservative 45; Mismatches 83; Indels 0; Gaps 0;

2 DTTVPFSLAELOOGLHODEFRRLRDKGLFYLTDCGLTDLTKSAKDLVIDFFEHGSEA 61

3 DSGIQIFDLDEHGVRLDSFRKSLFERGVYFREDSDSIKTEHAKAMDAMDVLFGNSAE 62

62 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITWQ 121

63 QKALRNLTNVRGFSDDLREASTARKGETDYSWYSIGLTDNLFSPAFALWTG 122

122 YFDRQYTSARAVAREVLRTGTEPDGGVAFDCEPLLRFPRFPQVPEHRSAREQPLRMA 181

123 YFDRFYARTQDIARSVLRASDAGLDESVDDELDCLLRLFRFPFPEVPEHRSAREQPLRMA 182

182 PHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVQKAP 241

183 PHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVQKAP 242

242 RHHVAAAPRRDQIAGSSRTSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGGN 301

243 RHQVAPSPMQORVGSSRTSVFFLRPKPDRFVSUPLARASGLDVLDFDGTATFGEWIGGN 302

302 YVNIR 306

303 YVNLR 307

RESULT 7

A29711

deacetoxycephalosporin C synthetase - fungus (Acremonium sp.)

Alternate names: expandase; hydroxylase

Species: Acremonium sp.

Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 19-May-2000

C:Accession: A29711; A41864

R:Samson, S.M.; Dotzlat, J.E.; Slisz, M.L.; Becker, G.W.; Van Frank, R.M.; Veal, L.E.; Ye

Bio/Technology 5, 1207-1214, 1987

A>Title: Cloning and expression of the fungal expandase/hydroxylase gene involved in cepi

A:Reference number: A29711

A:Accession: A29711

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-332 <SAM>

A>Note: the source is designated as Cephalosporium acremonium

R:Gutierrez, S.; Velasco, J.; Fernandez, F.J.; Martin, J.F.

J. Bacteriol. 174, 3056-3064, 1992

A>Title: The cefG gene of Cephalosporium acremonium is linked to the cefEF gene and encod

A:Reference number: A41864; MUID:92234966; PMID:1569032

A:Accession: A41864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-76, 'L' <GUT>

A:Experimental source: strain C10

A>Note: sequence extracted from NCBI backbone (NCBI:104773, NCBI:97574); this ORF is nc

A>Note: the source is designated as Cephalosporium acremonium

C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 56.6%; Score 926.5; DB 2; Length 332;

Best Local Similarity 56.7%; Pred. No. 2.9e-76;

Matches 177; Conservative 43; Mismatches 91; Indels 1; Gaps 1;

1 MDTTVPFSLAELOOGLHODEFRRLRDKGLFYLTDCGLTDLTKSAKDLVIDFFEHGSE 60

1 MTSKVPFRLDOLSKGKVLTELAAVTKGIFYLTESGLVDHDSARETCVDFPKNGSE 60

61 EKRAVTSVPVPMRRGFTGLESESTAQITNTGSDYSCYSGMTADNLFPSPGDFGRITW 120

61 EKRAVTLADRNRGFSALEWESTAVVTETGKYSYDSTCYSMGIGGNLFPNRPEDVMQ 120

121 QYFDRQYTSARAVAREVLRTGTEPDG-GVEAFDCEPLLRFPRFPQVPEHRSAREQPLR 179

121 DYFDRMYGAADVARAVLNSVGAFLAGEDIDDFVECDPLRLRYPEVPEHRSAREQPLR 180

180 MAPHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVQK 239

181 MGPHYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVQK 240

240 APRHHVAAAPRRDQIAGSSRTSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIG 299

241 APGRHVKSFGDRQVGSRTSVFFLRPKPDRFVSUPLARASGLDVLDFDGTATFGEWIG 300

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

[illegible]

THIS PAGE BLANK (USPTO)

```

RESULT 11
:08218
:isopenicillin N synthase (EC 1.14.11.-) pcbC [similarity] - Flavobacterium sp. (strain S
:Species: Flavobacterium sp.
:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 26-May-2000
:Accession: S08218
:Shiffman, D.; Cohen, G.; Aharonowitz, Y.; Palissa, H.; von Dohren, H.; Kleinkauf, H.;
:Kleinkauf, H.;
:Title: Nucleotide sequence of the isopenicillin N synthase gene (pcbC) of the gram neg
:Reference number: S08218; MUID:90174998; PMID:2308852
:Accession: S08218
:Molecule type: DNA
:Residues: 1-326 <SHI>
:Cross-references: EMBL:X17355; NID:g43423; PIDN:CAA35233.1; PID:g43424
:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTELKSAKDLVIDDFEHGSEAEK-----RAVTSPTVPTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87

iy 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPS-----GDFGRIW 119
ib 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDARHPGMRFF 139

iy 120 TOYFDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPQVPE 169
ib 140 EAYFSDVFDVAAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRIYPFLENYP- 193

iy 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
ib 194 -----PLKLGPDGKLSFEHHQDVSLITVLYQTAPN---LQVETAEGLDIPVSDE 242

iy 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
ib 243 HFLVNCGYMAHITNGYYPAPVHRV-----KYINAERLSIPFF 280

:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTELKSAKDLVIDDFEHGSEAEK-----RAVTSPTVPTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87

iy 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPS-----GDFGRIW 119
ib 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDARHPGMRFF 139

iy 120 TOYFDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPQVPE 169
ib 140 EAYFSDVFDVAAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRIYPFLENYP- 193

iy 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
ib 194 -----PLKLGPDGKLSFEHHQDVSLITVLYQTAPN---LQVETAEGLDIPVSDE 242

iy 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
ib 243 HFLVNCGYMAHITNGYYPAPVHRV-----KYINAERLSIPFF 280

:Genetics:
:Gene: pcbC
:Superfamily: isopenicillin N synthase
:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase
:45,209,265/Binding site: iron (His) #status predicted

Query Match      9.9%; Score 162; DB 2; Length 326;
Best Local Similarity 22.9%; Pred. No. 6.6e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

iy 24 RCLRDKGLFLYDCGLTDTLTELKSAKDLVIDDFEHGSEAEK-----RAVTSPTVPTMRRGFT 78
ib 32 RACRGSGFFYAANGV---DLAALQKFTTDWHMAMSAEEKWELAIRAYNPANPRNRNGY- 87

```

```

QY 79 GLESESTAQITNTGSDYSCYMGMTAD-----NLFPS-----GDFGRIW 119
DB 88 -----YMAVEGKKANESFCYLNPSFDADHATIKAGLPSHEVNIWPDARHPGMRFF 139
QY 120 TOYFDRQYTSRAVAREVLRAATGTEPDGGVEAF-----LDCEPLLRFRYPQVPE 169
DB 140 EAYFSDVFDVAAILRGFAIALGRE-----ESFFERHFSMDDTLSAVSLRIYPFLENYP- 193
QY 170 HRSABEQLRMAP-----HYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPD 220
DB 194 -----PLKLGPDGKLSFEHHQDVSLITVLYQTAPN---LQVETAEGLDIPVSDE 242
QY 221 AVLVFCAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSVFF 264
DB 243 HFLVNCGYMAHITNGYYPAPVHRV-----KYINAERLSIPFF 280

RESULT 13
DB4713
probable dioxigenase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84713
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84713
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <STO>
A:Cross-references: GB:AE002093; NID:g3201612; PIDN:AAC20719.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg30840
A:Map position: 2
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match      8.9%; Score 146.5; DB 2; Length 362;
Best Local Similarity 23.2%; Pred. No. 1.9e-05;
Matches 68; Conservative 40; Mismatches 118; Indels 67; Gaps 11;

QY 5 VPTFSLAELOQLHQDEF-----RRCLRDKGLFLYDCGLTDTLTELKSAKDLVID 53
DB 57 IPTIDL-----KGVFDEYTVTRESVIAMIRDAVERFCFFQVINHGINSNDVMEKMGDGRG 112
QY 54 FFEHSGAEKRAVTSPTVPTMRRGFTGLESESTAQITNTGSDYSCYMGMTAD-----N 108
DB 113 FHEQSDVRKKFY-----TRDVTKTVKYNSNFDLYSSPSANNRDTLS 154
QY 109 LFPSGDF-----GRITWQYFDRQYTSRAVAREVLRAATGTEPDGGVEAFDCEP- 157
DB 155 CFMAPDVPETEDLPDICEIMLEAKVMKLGELIFELLSALGNLPHLKE--MDCTKG 212
QY 158 -LLAFRYPQVPEHRSABEQLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLP 216
DB 213 LLMLSHYPPCPPE-----PGLTFTGSPHSRDSRFLILQ----DHIGGLQVQNGVWDVP 264
QY 217 YRPDAVLVFCGAIATLVGGQVAPRHVAAAPRRDQIAGSSRTSSV--FFLRP 267
DB 265 PVPGALLVNLGDLQLMTNDQFVSVEHRVLANK-----GEKPRISVASFFVHP 312

RESULT 14
DB6635
probable gibberellin 20-oxidase T7P1.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96635
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.;

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:57 ; Search time 42 Seconds
(without alignments)
1175.332 Million cell upd.

Title: US-09-582-486-1

Perfect score:

Sequence: 1 MDTTVPTFSLAELOQLHQD.....ATFQDWIGNVNIRRTSKA 311

Scoring table: BLOSUM62

scoring code: 1200002
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs. 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Minimum DB seq length:	0
Maximum DB seq length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1980.DAT.*
2:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1981.DAT.*
3:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1982.DAT.*
4:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1983.DAT.*
5:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1984.DAT.*
6:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1985.DAT.*
7:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1986.DAT.*
8:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1987.DAT.*
9:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1988.DAT.*
10:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1989.DAT.*
11:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1990.DAT.*
12:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1991.DAT.*
13:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1992.DAT.*
14:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1993.DAT.*
15:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1994.DAT.*
16:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1995.DAT.*
17:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1996.DAT.*
18:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1997.DAT.*
19:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1998.DAT.*
20:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA1999.DAT.*
21:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA2000.DAT.*
22:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA2001.DAT.*
23:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA2002.DAT.*
24:	/SIDSI/cgcdata/Geneseq/Geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1637	100.0	311	20	AAV22220	Deacetoxycephalexop
2	1629	99.5	311	18	AAW17775	S. clavuligerus wi
3	1629	99.5	311	23	AAU1044	Streptomyces clav
4	1627	99.4	311	23	AAU1050	Streptomyces clav
5	1626	99.3	311	23	AAU1046	Streptomyces clav
6	1626	99.3	311	23	AAU1051	Streptomyces clav
7	1625	99.3	310	23	AAU1053	S. clavuligerus pe
8	1625	99.3	310	23	AAU1525	S. clavuligerus pe
9	1625	99.3	311	18	AAW34168	Expandase mutant R

10	1625	99.3	311	18	AAW34176	Expandase mutant R
11	1624	99.2	311	18	AAW34175	Expandase mutant R
12	1624	99.2	311	22	AAE12016	Streptomyces clav
13	1624	99.2	311	23	AAU11047	Streptomyces clav
14	1624	99.2	311	24	ABG73769	S. clavuligerus mo
15	1623	99.1	311	18	AAW34167	Expandase mutant R
16	1623	99.1	311	18	AAW34167	Expandase mutant R
17	1623	99.1	311	23	AAU11048	Streptomyces clav
18	1622	99.1	310	23	AAU11528	S. clavuligerus pe
19	1622	99.1	311	23	AAU11052	Streptomyces clav
20	1621	99.0	311	10	ADP93215	Deactoxycyphalosp
21	1621	99.0	311	18	AAW34166	Expandase mutant R
22	1621	99.0	311	18	AAW34169	Expandase mutant R
23	1621	99.0	311	18	AAW34174	Expandase mutant P
24	1621	99.0	311	19	AAW40424	S. clavuligerus ex
25	1621	99.0	311	23	AAU11049	Streptomyces clav
26	1620	99.0	309	23	AAU11054	S. clavuligerus pe
27	1620	99.0	309	23	AAU11526	S. clavuligerus pe
28	1620	99.0	310	23	AAU11531	S. clavuligerus pe
29	1620	99.0	310	23	AAU11532	S. clavuligerus pe
30	1620	99.0	310	23	AAU11533	S. clavuligerus pe
31	1620	99.0	310	23	AAU11534	S. clavuligerus pe
32	1620	99.0	310	23	AAU11535	S. clavuligerus pe
33	1620	99.0	311	18	AAW34171	Expandase mutant C
34	1620	99.0	311	18	AAW34173	Expandase mutant P
35	1619	98.9	311	18	AAW34170	Expandase mutant C
36	1619	98.9	311	18	AAW34172	Expandase mutant C
37	1617	98.8	309	23	AAU11529	S. clavuligerus pe
38	1616	98.7	308	23	AAU11521	S. clavuligerus pe
39	1616	98.7	308	23	AAU11527	S. clavuligerus pe
40	1616	98.7	311	19	AAW40425	S. clavuligerus ex
41	1615	98.7	311	19	AAW40426	S. clavuligerus ex
42	1614	98.6	311	19	AAW40428	S. clavuligerus ex
43	1614	98.6	311	19	AAW40429	S. clavuligerus ex
44	1613	98.5	308	23	AAU11530	S. clavuligerus pe
45	1612	98.5	311	19	AAW40427	S. clavuligerus ex

ALIGNMENTS

XX WPI; 1999-405515/34.
 XX Three-dimensional crystal structures of deacetoxycephalosporin C
 PT synthase useful for the production of beta-lactams
 XX
 PS Claim 1; Page 11; 178pp; English.
 XX
 CC The invention relates to three-dimensional crystal structures of
 CC deacetoxycephalosporin C (DAOCC) synthase (DAOCS), which has this amino
 CC acid sequence. Micro-organisms capable of expressing a gene encoding a
 CC modified DAOCS are useful for the production of beta-lactams of the
 CC penicillin or cephalosporin (including cepham) families. The 3D structure
 CC of DAOCS is used to determine or predict the structure of another related
 CC 2-oxoglutarate dependent enzyme (such as deacetylcephalosporin C synthase
 CC (DAOCS3)) or related enzymes not from the penicillin and cephalosporin
 CC biosynthesis pathway. The structural data is used to modify the other
 CC enzyme or for designing an inhibitor for the other enzyme. The other
 CC enzyme may be prolyl 4/3 or aspartyl/lysyl-hydroxylase and the inhibitor
 CC is used for the treatment of human disease including fibrotic diseases
 CC including liver cirrhosis and arthritis. The other enzyme may be
 CC p-hydroxyphenylpyruvate dioxygenase and the inhibitor used to treat
 CC certain genetic disorders. Alternatively, the other enzyme is involved in
 CC herbicide resistance and the information is used to design new herbicides
 CC to overcome the problem of resistance. The crystal structure data allows
 CC modification of the enzyme to accept (or accept more efficiently)
 CC unnatural penicillin substrates for the preparation of new or
 CC commercially valuable antibacterial materials, or to enable the modified
 CC enzyme to produce unnatural (e.g. exomethylene cepham) or optimise the
 CC production of minor products (e.g. 3-beta-hydroxycephams) for use as
 CC antibacterials or as intermediates in the preparation of antibacterials
 CC or commercially valuable compounds.
 XX
 XX Sequence 311 AA;

Query Match 100.0%; Score 1637; DB 20; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.1e-172;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDTVTFTSLAELOOGLHODEFRRLDKGLFYLTDGLTDTLTKSAKDLVIDFFHSGE 60
 DB 1 MDTVTFTSLAELOOGLHODEFRRLDKGLFYLTDGLTDTLTKSAKDLVIDFFHSGE 60
 QY 61 AEKRAVTSVPVPTWRRGFTGLSESTAQITNTGYSYDSCYSMTADNLPSPGDFGRITW 120
 DB 61 AEKRAVTSVPVPTWRRGFTGLSESTAQITNTGYSYDSCYSMTADNLPSPGDFGRITW 120
 QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVCGAFTDLPRPDVAVFCGATATLVGGQVKA 240
 DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVCGAFTDLPRPDVAVFCGATATLVGGQVKA 240
 QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSKA 311
 DB 301 NYVNIIRRTSKA 311

RESULT 2
 AAW17775
 ID AAW17775 standard; protein; 311 AA.

XX AAW17775;
 AC AAW17775;
 DT 17-FEB-1998 (first entry)
 XX
 DE S. clavuligerus wild type expandase.

XX Streptomyces clavuligerus.
 KW Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
 KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
 KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
 KW cephalosporin ring; DOAC.
 XX
 OS Streptomyces clavuligerus.
 PN W09720053-A2.
 XX
 PD 05-JUN-1997.
 XX
 PF 27-NOV-1996; 96WO-BP05358.
 XX
 PR 27-NOV-1995; 95US-0007564.
 PR 27-NOV-1995; 95EP-0203259.
 XX
 PA (KONN) GIST-BROCADES BV.
 PA (UYOX-) UNIV OXFORD.
 XX
 PI Bovenberg RAL, Sutherland JD, Van Der Laan JM;
 XX
 DR WPI; 1997-310608/28.
 XX
 PT Expandase enzymes with mutation(s) to alter substrate specificity -
 PT used particularly for production of
 PT 7-amino-desacetoxycephalosporanic acid
 XX
 PS Disclosure; Fig 1; 21pp; English.
 CC This sequence represents the wild type expandase from Streptomyces
 CC clavuligerus. The DNA encoding this sequence was subjected to site
 CC directed mutagenesis to create the mutant expandase sequences of the
 CC invention (see AAW34166-W34177). The mutant expandase sequences have an
 CC altered substrate specificity compared to the wild type expandase.
 CC Expandase catalyses the expansion of the 5-membered thiazolidine ring of
 CC penicillin N to the 6-membered dihydrothiazine ring of DOAC. The mutant
 CC expandase gene is used particularly for the production of
 CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
 CC intermediates that is used in the preparation of semi-synthetic
 CC cephalosporins. The mutant expandases have improved ability to expand
 CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
 CC other penicillins such as penicillin N and isopenicillin N can act as
 CC competing substrates.
 XX
 SQ Sequence 311 AA;
 Query Match 99.5%; Score 1629; DB 18; Length 311;
 Best Local Similarity 99.7%; Pred. No. 8.5e-172;
 Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDTVTFTSLAELOOGLHODEFRRLDKGLFYLTDGLTDTLTKSAKDLVIDFFHSGE 60
 DB 1 MDTVTFTSLAELOOGLHODEFRRLDKGLFYLTDGLTDTLTKSAKDLVIDFFHSGE 60
 QY 61 AEKRAVTSVPVPTWRRGFTGLSESTAQITNTGYSYDSCYSMTADNLPSPGDFGRITW 120
 DB 61 AEKRAVTSVPVPTWRRGFTGLSESTAQITNTGYSYDSCYSMTADNLPSPGDFGRITW 120
 QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 DB 121 QYFDRQYTASRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVCGAFTDLPRPDVAVFCGATATLVGGQVKA 240
 DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVCGAFTDLPRPDVAVFCGATATLVGGQVKA 240
 QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSKA 311
 DB 301 NYVNIIRRTSKA 311

Db 301 NYNIRRTSKA 311

RESULT 3
AAU11044

ID AAU11044 standard; Protein; 311 AA.
AC AAU11044;
XX
DT 12-MAR-2002 (first entry)
XX
DE Streptomyces clavuligerus penicillin N expandase protein sequence.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid;
KW cephalosporin; cephalixin; penicillin V.
XX
OS Streptomyces clavuligerus.
XX
PN WO200185951-A1.
XX
PD 15-NOV-2001.
XX
PF 09-MAY-2001; 2001WO-GB02047.
XX
PR 09-MAY-2000; 2000GB-0011185.
XX
PA (ACSD-) ACS DOBFAR UK LTD.
XX
PI Johnson RI, Newbert RW;
XX
DR WPI; 2002-075247/10.
DR N-PSDB; AAS17241.
XX
XX New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxycephalosporanic acid, comprises
PT penicillin expandase having increased specificity for substrates such
PT as penicillin G -
XX
PS Claim 4; Fig 1; 42pp; English.
XX
CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxycephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC penicillin G. The present protein sequence represents the penicillin N
CC expandase enzyme of the invention.
XX
SQ Sequence 311 AA;

Query Match 99.5%; Score 1629; DB 23; Length 311;
Best Local Similarity 99.7%; Pred. No. 8.5e-172;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAEQQQLHQDEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVDPFRHGS 60
Db 1 MDTTPTFSLAEQQQLHQDEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVDPFRHGS 60

QY 61 AEKRVTSPTVPTWRGFTGLESESTAQITNTGYSYDSMYSMTADNLPSPGDFGRIT 120
Db 61 AEKRVTSPTVPTWRGFTGLESESTAQITNTGYSYDSMYSMTADNLPSPGDFGRIT 120

QY 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFIDCEPLLRFYFPQVPEHRSABEOLPM 180
|||||

Db 121 QYFDRQYTASRAVAREVLRTATGTEPDGGVEAFIDCEPLLRFYFPQVPEHRSABEOLPM 180

QY 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGGAFDTLPRPDVAVLVFCGAIATLVTTGGQVKA 240
|||||

Db 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGGAFDTLPRPDVAVLVFCGAIATLVTTGGQVKA 240

QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIGG 300
|||||

Db 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYNIRRTSKA 311
|||||

Db 301 NYNIRRTSKA 311

RESULT 4
AAU11050

ID AAU11050 standard; Protein; 311 AA.
XX
AC AAU11050;
XX
DT 12-MAR-2002 (first entry)
XX
DE Streptomyces clavuligerus penicillin N expandase I305L mutant.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid;
KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 305 /note= "Wild-type Ile substituted by Leu"
FT
XX
PN WO200185951-A1.
XX
PD 15-NOV-2001.
XX
PF 09-MAY-2001; 2001WO-GB02047.
XX
PR 09-MAY-2000; 2000GB-0011185.
XX
PA (ACSD-) ACS DOBFAR UK LTD.
XX
PI Johnson RI, Newbert RW;
XX
DR WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxycephalosporanic acid, comprises
PT penicillin expandase having increased specificity for substrates such
PT as penicillin G -
XX
PS Claim 8; Page -; 42pp; English.
XX
CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxycephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC penicillin G. This sequence is one of a collection of Streptomyces
CC clavuligerus mutants, AAU11046-AAU11054 and AAU11521-AAU11537, used in
CC the invention to improve the ring-expanding activity of expandase.

CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU11044).

XX SQ Sequence 311 AA;

Query Match 99.4%; Score 1627; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.4e-171;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAELOQGLHODEFRRLDKGLFYLTCGLTDTLTKSAKDLVDFEHSSE 60
DB 1 MDTTPTFSLAELOQGLHODEFRRLDKGLFYLTCGLTDTLTKSAKDLVDFEHSSE 60

QY 61 AEKRAVTSVPVTRRGFTGLESESTAIQNTGSDYSYSCYSGMTADNLPFGDGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAIQNTGSDYSYSCYSGMTADNLPFGDGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPPHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPPHRSAAEQPLRM 180

QY 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 5
AAU11046
ID AAU11046 standard; Protein; 311 AA.
AC AAU11046;
XX
DT 12-MAR-2002 (first entry)
XX Streptomyces clavuligerus penicillin N expandase L158V mutant.
DE Penicillin N expandase; ring-expanding; penicillin G;
XX phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW cephalosporin; cephalaxin; penicillin V; mutant; mutein.
XX Streptomyces clavuligerus.
OS Synthetic.
XX Key Location/Qualifiers

FT Misc-difference 158
FT /note= "Wild-type Leu substituted by Val"
XX
XX WO200185951-A1
XX
PD 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB02047.
XX
XX 09-MAY-2000; 2000GB-0011185.
XX
XX (ACSD-) ACS DOBFAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT penicillin expandase having increased specificity for substrates such

PT as penicillin G -
XX
PS Claim 6; Page -; 42pp; English.
XX
CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalaxin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin N expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC penicillin G. This sequence is one of a collection of Streptomyces
CC clavuligerus mutants, AAU11046-AAU11054 and AAU11521-AAU11537, used in
CC the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU11044).

XX SQ Sequence 311 AA;

Query Match 99.3%; Score 1626; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.8e-171;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAELOQGLHODEFRRLDKGLFYLTCGLTDTLTKSAKDLVDFEHSSE 60
DB 1 MDTTPTFSLAELOQGLHODEFRRLDKGLFYLTCGLTDTLTKSAKDLVDFEHSSE 60

QY 61 AEKRAVTSVPVTRRGFTGLESESTAIQNTGSDYSYSCYSGMTADNLPFGDGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLESESTAIQNTGSDYSYSCYSGMTADNLPFGDGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPPHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGVEAFDCEPLLRFRYPVPPHRSAAEQPLRM 180

QY 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMTVLIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIATLVGGQVKA 240

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 6
AAU11051
ID AAU11051 standard; Protein; 311 AA.
XX
AC AAU11051;
XX
DT 12-MAR-2002 (first entry)
XX Streptomyces clavuligerus penicillin N expandase I305M mutant.
DE Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
KW cephalosporin; cephalaxin; penicillin V; mutant; mutein.
XX Streptomyces clavuligerus.
OS Synthetic.
XX Key Location/Qualifiers

FT Misc-difference 158
FT /note= "Wild-type Leu substituted by Val"
XX
XX WO200185951-A1
XX
PD 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB02047.
XX
XX 09-MAY-2000; 2000GB-0011185.
XX
XX (ACSD-) ACS DOBFAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT penicillin expandase having increased specificity for substrates such

FT Misc-difference 305 /note= "Wild-type Ile substituted by Met"
FT XX WO200185951-A1.
FT XX 15-NOV-2001.
FT XX 09-MAY-2001; 2001WO-GB02047.
FT XX 09-MAY-2000; 2000GB-0011185.
FT XX (ACSD-) ACS DOBFAF UK LTD.
FT XX Johnson RI, Newbert RW;
FT XX WPI; 2002-075247/10.
FT XX New enzyme useful for ring-expanding penicillin G to produce
FT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
FT penicillin N expandase having increased specificity for substrates such
FT as penicillin G -
FT Claim 8; Page -: 42pp; English.
FT XX The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin N expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC clavuligerus mutants, AAU1046-AAU1054 and AAU1521-AAU1537, used in
CC penicillin G. This sequence is one of a collection of Streptomyces
CC the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU1044).
XX SQ Sequence 311 AA;
Query Match 99.3%; Score 1626; DB 23; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.8e-171;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTVPFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKOLVIDFFHGE 60
DB 1 MDTTVPFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKOLVIDFFHGE 60
QY 61 AEKRAVTSPTVPTMRGFTGLESESTAQITNTGYSYDYSYMGTSMTADNLPSPGDFGRIT 120
DB 61 AEKRAVTSPTVPTMRGFTGLESESTAQITNTGYSYDYSYMGTSMTADNLPSPGDFGRIT 120
QY 121 QYFDROYTASRAVAREVLRTATGTEPDGVEAFIDCEPFLFRFPQVPEHRSABEOLRM 180
DB 121 QYFDROYTASRAVAREVLRTATGTEPDGVEAFIDCEPFLFRFPQVPEHRSABEOLRM 180
QY 181 APHYDLSMWTLIOOTPCANGFVSLQAEVGGAFDLPVRPAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMWTLIOOTPCANGFVSLQAEVGGAFDLPVRPAVLVFCGAIATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSRSSVFFLRPNADFTFSVPLARECGFVSLDGTATFDQWIGG 300
DB 241 PRHHVAAPRRDQIAGSRSSVFFLRPNADFTFSVPLARECGFVSLDGTATFDQWIGG 300
QY 301 NYNVRTSKA 311
DB 301 NYNVRTSKA 311

RESULT 7
AAU1053
ID AAU1053 standard; Protein; 310 AA.
XX AC AAU1053;
XX DT 12-MAR-2002 (first entry)
XX DE S. clavuligerus penicillin N expandase 1 residue deletion at C-terminus.
XX KW Penicillin N expandase; ring-expanding; penicillin G;
XX KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
XX KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
XX OS Streptomyces clavuligerus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 310 /note= "Protein is truncated at this residue"
XX FT FT
XX PN WO200185951-A1.
XX PD 15-NOV-2001.
XX PF 09-MAY-2001; 2001WO-GB02047.
XX PR 09-MAY-2000; 2000GB-0011185.
XX PA (ACSD-) ACS DOBFAF UK LTD.
XX PI Johnson RI, Newbert RW;
XX WPI; 2002-075247/10.
XX PT New enzyme useful for ring-expanding penicillin G to produce
PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
PT penicillin N expandase having increased specificity for substrates such
PT as penicillin G -
XX Disclosure; Page -: 42pp; English.
XX The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not
CC the natural substrate of the unmodified expandase. The invention is
CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
CC as a starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as
CC a probe. The modification of the penicillin N expandase enhances the
CC activity of the enzyme such as penicillin N expandase for penicillin
CC G or V as a substrate. The modified expandase has enhanced catalytic
CC activity or increased specificity for another substrate such as
CC clavuligerus mutants, AAU1046-AAU1054 and AAU1521-AAU1537, used in
CC penicillin G. This sequence is one of a collection of Streptomyces
CC the invention to improve the ring-expanding activity of expandase.
CC Note: The present protein sequence is not shown in the specification
CC but is derived from the Streptomyces clavuligerus wild-type
CC penicillin N expandase sequence given in figure 1 (see AAU1044).
XX SQ Sequence 310 AA;
Query Match 99.3%; Score 1625; DB 23; Length 310;
Best Local Similarity 99.7%; Pred. No. 2.3e-171;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTVPFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKOLVIDFFHGE 60
DB 1 MDTTVPFSLAELOOGLHQDEFRCLRDKGLFYLTCGLTDTLTKSAKOLVIDFFHGE 60

QY 61 AEKRAVTSVPPTWRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120
 DB 61 AEKRAVTSVPPTWRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120
 QY 121 QYFDRQYTSRAVAREVLRTATGTPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 DB 121 QYFDRQYTSRAVAREVLRTATGTPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
 QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRPPDAVLVFCGAIATLVGGQVKA 240
 DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRPPDAVLVFCGAIATLVGGQVKA 240
 QY 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSK 310
 DB 301 NYVNIIRRTSK 310
 RESULT 8
 AAU11525
 ID AAU11525 standard; Protein; 310 AA.
 AC AAU11525;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE S. clavuligerus penicillin N expandase K310 deletion mutant.
 XX
 KW Penicillin N expandase; ring-expanding; penicillin G;
 KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
 KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
 XX
 OS Streptomycetes clavuligerus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 310 /note= "Protein is truncated at this residue"
 FT
 XX WO200185951-A1.
 XX
 PD 15-NOV-2001.
 XX
 XX 09-MAY-2001; 2001WO-GB02047.
 XX
 XX 09-MAY-2000; 2000GB-0011185.
 XX
 XX (ACSD-) ACS DOBFAR UK LTD.
 PA
 XX Johnson RI, Newbert RW;
 XX
 XX WPI; 2002-075247/10.
 XX
 XX New enzyme useful for ring-expanding penicillin G to produce
 PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
 PT penicillin expandase having increased specificity for substrates such
 PT as penicillin G -
 XX
 XX Examples; Page -; 42pp; English.
 XX
 CC The present invention relates to new penicillin expandases modified to
 CC improve the ring-expanding activity for a substrate which is not
 CC the natural substrate of the unmodified expandase. The invention is
 CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
 CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
 CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
 CC as a starting point for the production of a range of semi-synthetic
 CC cephalosporins, most notably cephalixin. The polynucleotide of the
 CC invention is useful for producing a primer e.g. a PCR primer and as

CC a probe. The modification of the penicillin expandase enhances the
 CC activity of the enzyme such as penicillin N expandase for penicillin
 CC G or V as a substrate. The modified expandase has enhanced catalytic
 CC activity or increased specificity for another substrate such as
 CC penicillin G. This sequence is one of a collection of Streptomycetes
 CC clavuligerus mutants, AAU11046-AAU11054 and AAU11521-AAU11537, used in
 CC the invention to improve the ring-expanding activity of expandase.
 CC Note: The present protein sequence is not shown in the specification
 CC but is derived from the Streptomycetes clavuligerus wild-type
 CC penicillin N expandase sequence given in figure 1 (see AAU11044).

XX SQ Sequence 310 AA;

Query Match 99.3%; Score 1625; DB 23; Length 310;

Best Local Similarity 99.7%; Pred. No. 2.3e-171;

Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTVPFTSLAELOQLQHODEFRRCRLDKGLFVLTDCGLTDTLKSADLVIDFFEHGSE 60

DB 1 MDTTVPFTSLAELOQLQHODEFRRCRLDKGLFVLTDCGLTDTLKSADLVIDFFEHGSE 60

QY 61 AEKRAVTSVPPTWRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120

DB 61 AEKRAVTSVPPTWRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPSPGDFGRITW 120

QY 121 QYFDRQYTSRAVAREVLRTATGTPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

DB 121 QYFDRQYTSRAVAREVLRTATGTPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRPPDAVLVFCGAIATLVGGQVKA 240

DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRPPDAVLVFCGAIATLVGGQVKA 240

QY 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300

DB 241 PRHHVAAPRRDQTAGSSRTSSVFFLRPNADFTFVPLARECGFDVSLDGETATFQDWIGG 300

QY 301 NYVNIIRRTSK 310

DB 301 NYVNIIRRTSK 310

RESULT 9

AAW34168

ID AAW34168 standard; protein; 311 AA.

XX AAW34168;

AC AAW34168;

XX 17-FEB-1998 (first entry)

DT 17-FEB-1998 (first entry)

XX Expandase mutant R74Q.

DE Expandase mutant R74Q.

XX Expandase; 7-aminodesacetoxyccephalosporanic acid; 7-ADCA production;

KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;

KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;

KW cephalosporin ring; DOAC; mutant.

XX Streptomycetes clavuligerus.

OS Streptomyces

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 74 /note= "Arg to Gln mutation"

FT

XX WO9720053-A2.

PN

XX 05-JUN-1997.

PD

XX 27-NOV-1996; 96WO-EP05358.

PF

XX 27-NOV-1995; 95US-0007564.

XX 27-NOV-1995; 95EP-0203259.

PR

XX

PA	(KONN) GIST-BROCADES BV.	KW	cephalosporin ring; DOAC; mutant.
PA	(UYOX-) UNIV OXFORD.	XX	
PI	Bovenberg RAL, Sutherland JD, Van Der Laan JM;	OS	Streptomyces clavuligerus.
XX		OS	Synthetic.
DR	WPI; 1997-310608/28.	PH	Key
XX		FT	Location/Qualifiers
PT	Expandase enzymes with mutation(s) to alter substrate specificity -	FT	Misc-difference 266
PT	used particularly for production of	XX	/note= "Arg to Gln mutation"
PT	7-amino:desacetoxy:cephalosporanic acid	XX	
XX		PN	W09720053-A2.
PS	Claim 2; Page -: 21pp; English.	PD	05-JUN-1997.
XX		XX	
CC	AAW34166-W34177 represent mutant expandase sequences of the invention.	XX	
CC	These sequences represent mutated versions of the Streptomyces	PF	27-NOV-1996; 96WO-EP05358.
CC	clavuligerus expandase sequences shown in AAW17775. The DNA encoding the	PR	27-NOV-1995; 95US-0007564.
CC	wild type expandase was subjected to site directed mutagenesis to create	PR	27-NOV-1995; 95EP-0203259.
CC	these sequences. The mutant expandase sequences have an altered	XX	
CC	substrate specificity compared to the wild type expandase. Expandase	PA	(KONN) GIST-BROCADES BV.
CC	catalyses the expansion of the 5-membered thiazolidine ring of penicillin	PA	(UYOX-) UNIV OXFORD.
CC	N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase	XX	
CC	*gene is used particularly for the production of	PI	Bovenberg RAL, Sutherland JD, Van Der Laan JM;
CC	7-amino:desacetoxy:cephalosporanic acid (7-ADCA). 7-ADCA is one of the key	XX	
CC	intermediates that is used in the preparation of semi-synthetic	XX	
CC	cephalosporins. These sequences have improved ability to expand	DR	WPI; 1997-310608/28.
CC	penicillin G into a cephalosporin ring in vitro and/or in vivo where	XX	
CC	other penicillins such as penicillin N and isopenicillin N can act as	PT	Expandase enzymes with mutation(s) to alter substrate specificity -
CC	competing substrates.	PT	used particularly for production of
XX		PT	7-amino:desacetoxy:cephalosporanic acid
SQ	Sequence 311 AA;	XX	
		PS	Claim 2; Page -: 21pp; English.
		XX	
		CC	AAW34166-W34177 represent mutant expandase sequences of the invention.
		CC	These sequences represent mutated versions of the Streptomyces
		CC	clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
		CC	wild type expandase was subjected to site directed mutagenesis to create
		CC	these sequences. The mutant expandase sequences have an altered
		CC	substrate specificity compared to the wild type expandase. Expandase
		CC	catalyses the expansion of the 5-membered thiazolidine ring of penicillin
		CC	N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase
		CC	*gene is used particularly for the production of
		CC	7-amino:desacetoxy:cephalosporanic acid (7-ADCA). 7-ADCA is one of the key
		CC	intermediates that is used in the preparation of semi-synthetic
		CC	cephalosporins. These sequences have improved ability to expand
		CC	penicillin G into a cephalosporin ring in vitro and/or in vivo where
		CC	other penicillins such as penicillin N and isopenicillin N can act as
		CC	competing substrates.
		XX	
		SQ	Sequence 311 AA;
		Query Match	99.3%; Score 1625; DB 18; Length 311;
		Best Local Similarity	99.4%; Pred. No. 2.4e-171;
		Matches 309; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
QY	1	MDTTVTFTSLAEIQQGLHQDEFRRCRLDKGLFYLTCGLTDTLTKSAKOLVIDFFEHGSE	60
DB	1	MDTTVTFTSLAEIQQGLHQDEFRRCRLDKGLFYLTCGLTDTLTKSAKOLVIDFFEHGSE	60
QY	61	AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSYSDYSMCYSMGTDNLPPSGDGRIT	120
DB	61	AEKRAVTSVPVPTMRGFTGLESESTAQITNTGSYSDYSMCYSMGTDNLPPSGDGRIT	120
QY	121	QYFDROYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRFYFPQVPEHRSAAEQPLRM	180
DB	121	QYFDROYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRFYFPQVPEHRSAAEQPLRM	180
QY	241	PRHHVAAPRDOJAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGTATFQDWIGG	300
DB	241	PRHHVAAPRDOJAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGTATFQDWIGG	300
QY	301	NYVNIIRRTSKA 311	
DB	301	NYVNIIRRTSKA 311	
		RESULT 10	
		AAW34176	
ID	AAW34176	standard; protein; 311 AA.	
XX			
AC	AAW34176;		
XX			
DT	17-FEB-1998	(first entry)	
XX			
DE	Expandase mutant R2660.		
XX			
KW	Expandase; 7-amino:desacetoxy:cephalosporanic acid; 7-ADCA production;		
KW	thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;		
KW	semi-synthetic cephalosporin production; penicillin G; isopenicillin N;		

QY 301 NYVNIIRRTSKA 311
 Db 301 NYVNIIRRTSKA 311

RESULT 11
 AAW34175
 ID AAW34175 standard; protein; 311 AA.
 XX
 AC AAW34175;
 XX
 DT 17-FEB-1998 (first entry)
 XX
 DE Expandase mutant R266N.
 XX
 KW Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
 KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
 KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
 KW cephalosporin ring; DOAC; mutant.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 266 /note= "Arg to Asn mutation"
 FT
 PN W09720053-A2.
 XX
 XX 05-JUN-1997.
 XX
 XX 27-NOV-1996; 96WO-EP05358.
 PF
 XX 27-NOV-1995; 95US-0007564.
 PR
 PR 27-NOV-1995; 95EP-0203259.
 XX
 XX (KONN) GIST-BROCADES BV.
 PA (UYOX-) UNIV OXFORD.
 PA
 PI Bovenberg RAL, Sutherland JD, Van Der Laan JM;
 XX
 XX WPI; 1997-310608/28.
 DR
 XX
 XX Expandase enzymes with mutation(s) to alter substrate specificity -
 PT used particularly for production of
 PT 7-amino:desacetoxy:cephalosporanic acid
 XX
 XX +Claim 2; Page -; 2ipp; English.
 PS
 XX AAW34166-W34177 represent mutant expandase sequences of the invention.
 CC These sequences represent mutated versions of the Streptomyces
 CC clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
 CC wild type expandase was subjected to site directed mutagenesis to create
 CC these sequences. The mutant expandase sequences have an altered
 CC substrate specificity compared to the wild type expandase. Expandase
 CC catalyses the expansion of the 5-membered thiazolidine ring of penicillin
 CC N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase
 CC gene is used particularly for the production of
 CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
 CC intermediates that is used in the preparation of semi-synthetic
 CC cephalosporins. These sequences have improved ability to expand
 CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
 CC other penicillins such as penicillin N and isopenicillin N can act as
 CC competing substrates.
 XX
 SQ Sequence 311 AA;
 Query Match 99.2%; Score 1624; DB 18; Length 311;
 Best Local Similarity 99.4%; Pred. No. 3e-171;
 Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAELOQGLHQBDFRRCRLDKGLFYLTDCGLTDTTELKSAKDVLIDFFEHGSE 60
 |||||

Db 1 MDTTPTFSLAELOQGLHQBDFRRCRLDKGLFYLTDCGLTDTTELKSAKDVLIDFFEHGSE 60
 QY 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGYSYSDYSNMGYMGADNLFPSSGDFGRWT 120
 |||||
 Db 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGYSYSDYSNMGYMGADNLFPSSGDFGRWT 120
 |||||
 QY 121 QYFDRQYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYFFQVPEHRSABEQPLRM 180
 |||||
 Db 121 QYFDRQYTASRAVAREVLRTGTEPDGGVEAFDCEPLLRFRYFFQVPEHRSABEQPLRM 180
 |||||
 QY 181 APHYDLMSVTLTQQTFCANGFVSLQAEVGGAFDLPYRPDAVLVFCGATATLTGGQVKA 240
 |||||
 Db 181 APHYDLMSVTLTQQTFCANGFVSLQAEVGGAFDLPYRPDAVLVFCGATATLTGGQVKA 240
 |||||
 QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 |||||
 Db 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 |||||
 QY 301 NYVNIIRRTSKA 311
 |||||
 Db 301 NYVNIIRRTSKA 311
 |||||

RESULT 12
 AAE12016
 ID AAE12016 standard; Protein; 311 AA.
 XX
 AC AAE12016;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase.
 XX
 KW Deacetoxycephalosporin C synthetase; DAOCS; antibiotic; penicillin;
 KW oxygenase; non-haeme iron (II) dependent family; oxidase; mutant; muten.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 181 /note= "Wild type Ala substituted with Arg"
 FT
 PN US6284483-B1.
 XX
 XX 04-SEP-2001.
 PD
 PF 06-OCT-1999; 99US-0413231.
 XX
 XX 06-OCT-1999; 99US-0413231.
 PR
 XX (UNMS) UNIV MICHIGAN STATE.
 PA
 PI Dilley DR, Kadyrzhanova DK, Wang Z, Warner TW;
 XX WPI; 2001-615433/71.
 DR
 XX Method for producing antibiotics, particularly penicillin G or V,
 FT comprises employing a modified isopenicillin N synthetase, in either
 FT an organism or a cell-free system under the control of bicarbonate -
 XX
 PS Disclosure; Column 45-48; 27pp; English.
 XX
 CC The invention relates to a modified enzyme of a non-haeme iron (II)
 CC dependent family of oxygenases and oxidases which renders the enzyme
 CC dependent on bicarbonate for activity. The invention also related to a
 CC method for producing penicillin G or V comprises employing a modified
 CC enzyme, particularly isopenicillin N synthetase (IPNS), in either an
 CC organism or a cell-free system, under the control of bicarbonate. The
 CC modified enzymes are IPNS, deacetoxycephalosporin C synthetase (DAOCS)
 CC and deacetoxycephalosporin C synthetase (DACS). The method is useful for
 CC producing antibiotics, particularly penicillin G or V. The method is
 CC useful for making organisms useful for making an antibiotic dependent

CC on bicarbonate to make the antibiotic. The present sequence is
 CC Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase
 CC showing an Arg181 in place of Ala181.
 XX
 SQ Sequence 311 AA;
 Query Match 99.2%; Score 1624; DB 22; Length 311;
 Best Local Similarity 99.4%; Pred. No. 3e-171;
 Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDTTPTFSLAEIQQGLHQDEFRRCRLDKGLFYLTDGGLTDTLTKSAKDLVDFEFGSE 60
 DB 1 MDTTPTFSLAEIQQGLHQDEFRRCRLDKGLFYLTDGGLTDTLTKSAKDLVDFEFGSE 60
 QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIT 120
 DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIT 120
 QY 121 QYFDRQYTSRAVAREVLRATGTEPDGGVEAFDCEPDLFRFPQVPEHRSAEQPLRM 180
 DB 121 QYFDRQYTSRAVAREVLRATGTEPDGGVEAFDCEPDLFRFPQVPEHRSAEQPLRM 180
 QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGOVKA 240
 DB 181 RPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGOVKA 240
 QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSKA 311
 DB 301 NYVNIIRRTSKA 311
 RESULT 13
 AAU11047
 ID AAU11047 standard; Protein; 311 AA.
 XX
 AC AAU11047;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Streptomyces clavuligerus penicillin N expandase L158A mutant.
 XX
 KW Penicillin N expandase; ring-expanding; penicillin G;
 KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid;
 KW cephalosporin; cephalixin; penicillin V; mutant; mutein.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 158 /note= "Wild-type Leu substituted by Ala"
 FT
 FT W0200185951-Al.
 XX
 XX 15-NOV-2001.
 XX
 XX 09-MAY-2001; 2001WO-GB02047.
 XX
 XX 09-MAY-2000; 2000GB-0011185.
 XX
 XX (ACSD-) ACS DOBFAR UK LTD.
 XX
 XX Johnson RI, Newbert RW;
 XX
 XX WPI; 2002-075247/10.
 XX
 PT New enzyme useful for ring-expanding penicillin G to produce
 PT phenylacetyl-7-amino desacetoxyccephalosporanic acid, comprises
 PT penicillin expandase having increased specificity for substrates such

PT as penicillin G -
 XX
 PS Examples; Page -: 42pp; English.
 XX
 CC The present invention relates to new penicillin expandases modified to
 CC improve the ring-expanding activity for a substrate which is not
 CC the natural substrate of the unmodified expandase. The invention is
 CC useful for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA
 CC (amino desacetoxyccephalosporanic acid) and the phenylacetyl side chain
 CC is removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful
 CC as a starting point for the production of a range of semi-synthetic
 CC cephalosporins, most notably cephalixin. The polynucleotide of the
 CC invention is useful for producing a primer e.g. a PCR primer and as
 CC a probe. The modification of the penicillin expandase enhances the
 CC activity of the enzyme such as penicillin N expandase for penicillin
 CC G or V as a substrate. The modified expandase has enhanced catalytic
 CC activity or increased specificity for another substrate such as
 CC penicillin G. This sequence is one of a collection of Streptomyces
 CC clavuligerus mutants, AAU1046-AAU11054 and AAU1521-AAU1537, used in
 CC the invention to improve the ring-expanding activity of expandase.
 CC Note: The present protein sequence is not shown in the specification
 CC but is derived from the Streptomyces clavuligerus wild-type
 CC penicillin N expandase sequence given in figure 1 (see AAU11044).
 XX
 SQ Sequence 311 AA;
 Query Match 99.2%; Score 1624; DB 23; Length 311;
 Best Local Similarity 99.4%; Pred. No. 3e-171;
 Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MDTTPTFSLAEIQQGLHQDEFRRCRLDKGLFYLTDGGLTDTLTKSAKDLVDFEFGSE 60
 DB 1 MDTTPTFSLAEIQQGLHQDEFRRCRLDKGLFYLTDGGLTDTLTKSAKDLVDFEFGSE 60
 QY 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIT 120
 DB 61 AEKRAVTSVPVPTMRGFTGLESESTAQITNTGYSYDSCYSGMTADNLPSPGDFGRIT 120
 QY 121 QYFDRQYTSRAVAREVLRATGTEPDGGVEAFDCEPDLFRFPQVPEHRSAEQPLRM 180
 DB 121 QYFDRQYTSRAVAREVLRATGTEPDGGVEAFDCEPDLFRFPQVPEHRSAEQPLRM 180
 QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGOVKA 240
 DB 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGOVKA 240
 QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 DB 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
 QY 301 NYVNIIRRTSKA 311
 DB 301 NYVNIIRRTSKA 311
 RESULT 14
 ABG73769
 ID ABG73769 standard; Protein; 311 AA.
 XX
 AC ABG73769;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE S. clavuligerus modified DAOCS SEQ ID 18.
 XX
 KW Non-heme iron II dependent family; oxygenase; oxidase; antibiotic;
 KW iron ligand; enzyme; bicarbonate; penicillin G; penicillin V;
 KW cephalosporin C; isopenicillin N synthetase; IPNS; DAOCS;
 KW deacetoxyccephalosporin C synthetase.
 XX
 OS Streptomyces clavuligerus.
 OS Synthetic.
 XX

```
PN US2002127633-A1.
XX
PD 12-SEP-2002.
XX
XX 08-AUG-2001; 2001US-0924841.
XX PF
XX 06-OCT-1999; 99US-0413231.
XX PR
XX (UNMS ) UNIV MICHIGAN STATE.
XX PA
XX Dillely DR, Kadyrzhanova DK, Wang Z, Warner TW;
XX PI WPI; 2003-147393/14.
XX DR
XX Novel enzyme of non-heme iron II dependent family of oxygenases and
XX PT oxidases useful for producing antibiotic, has mutation that renders the
XX PT enzyme dependent on bicarbonate to produce the antibiotic
XX
XX Claim 5; Page 24-25; 31pp; English.
XX
XX This invention describes a novel enzyme of a non-heme iron II dependent
XX CC family of oxygenases and oxidases which is in a pathway to produce an
XX CC antibiotic. The enzyme comprises a mutation which is an amino acid
XX CC residue that is two amino acid residues upstream of a histidine residue
XX CC which is an iron ligand of the enzyme, where the mutation renders the
XX CC enzyme dependent on bicarbonate to produce the antibiotic. The enzyme
XX CC and modified organisms containing the enzyme are useful for producing an
XX CC antibiotic such as penicillin G or V and cephalosporin C. This sequence
XX CC represents an deacetoxycephalosporin C synthetase (DAOCS) variant
XX CC described in the disclosure of the invention.
XX
XX Sequence 311 AA;
XX
XX Query Match 99.2%; Score 1624; DB 24; Length 311;
XX Best Local Similarity 99.4%; Pred. No. 3e-171;
XX Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFEHGSE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFEHGSE 60
QY 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPFSGDGRITW 120
DB 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPFSGDGRITW 120
QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLLIQQTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
DB 181 RPHYDLSMVTLLIQQTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYNVIIRRTSKA 311
DB 301 NYNVIIRRTSKA 311
RESULT 15
AAW34167
ID AAW34167 standard; protein; 311 AA.
XX
XX AAW34167;
XX
XX 17-FEB-1998 (first entry)
XX DE
XX Expandase mutant R74M.
XX
XX Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KW thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KW semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KW cephalosporin ring; DOAC; mutant.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 74 /note= "Arg to Met mutation"
FT
XX
XX WO9720053-A2.
XX
XX 05-JUN-1997.
XX
XX 27-NOV-1996; 96WO-EP05358.
XX PF
XX 27-NOV-1995; 95US-0007564.
XX PR
XX 27-NOV-1995; 95EP-0203259.
XX PR
XX (KONN ) GIST-BROCADES BV.
XX PA (OYOK-) UNIV OXFORD.
XX
XX Bovenberg RAL, Sutherland JD, Van Der Laan JM;
XX PI WPI; 1997-310608/28.
XX
XX Expandase enzymes with mutation(s) to alter substrate specificity -
XX PT used particularly for production of
XX PT 7-amino:desacetoxy:cephalosporanic acid
XX
XX Claim 2; Page -; 21pp; English.
XX
XX AAW34166-W34177 represent mutant expandase sequences of the invention.
XX CC These sequences represent mutated versions of the Streptomyces
XX CC clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
XX CC wild type expandase was subjected to site directed mutagenesis to create
XX CC these sequences. The mutant expandase sequences have an altered
XX CC substrate specificity compared to the wild type expandase. Expandase
XX CC catalyses the expansion of the 5-membered thiazolidine ring of penicillin
XX CC N to the 6-membered dihydrothiazine ring of DOAC. The mutant expandase
XX CC gene is used particularly for the production of
XX CC 7-aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
XX CC intermediates that is used in the preparation of semi-synthetic
XX CC cephalosporins. These sequences have improved ability to expand
XX CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
XX CC other penicillins such as penicillin N and isopenicillin N can act as
XX CC competing substrates.
XX
XX Sequence 311 AA;
XX
XX Query Match 99.1%; Score 1623; DB 18; Length 311;
XX Best Local Similarity 99.4%; Pred. No. 3.9e-171;
XX Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFEHGSE 60
DB 1 MDTTPTFSLAELOOGLHODEFRRCRLDKGLFVLTDCGLTDTLTKSAKDLVIDFFEHGSE 60
QY 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPFSGDGRITW 120
DB 61 AEKRAVTSVPVPTMRRGFTGLESESTAQITNTGSDYSYSCYSGMTADNLPFSGDGRITW 120
QY 121 QYFDRQYTSRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSAAEQPLRM 180
DB 121 QYFDRQYTSRAVAREVLRTATGTEPDGGVEAFDCEPLLRFYFPQVPEHRSAAEQPLRM 180
QY 181 APHYDLSMVTLLIQQTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
DB 181 APHYDLSMVTLLIQQTPCANGFVLSQAEVGAFTDLPYRDAVLVFCGATATLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
```

Db 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIIRTSKA 311
Db 301 NYVNIIRTSKA 311

Search completed: February 3, 2004, 17:36:36
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:33:58 ; Search time 17 Seconds
(without alignments)
860.312 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTTPTFTSLAEIQQQLHQD.....ATFQDWIGGVNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1629	99.5	311	1	CEFE_STRCL
2	1186.5	72.5	314	1	CEFE_NOCCLA
3	965.5	59.0	310	1	CEFF_NOCCLA
4	939.5	57.4	318	1	CEFF_STRCL
5	926.5	56.6	332	1	EXPA_CEPAC
6	179	10.9	329	1	IPNS_STRJU
7	176	10.8	329	1	IPNS_STRGR
8	164	10.0	329	1	IPNS_STRCL
9	162	9.9	326	1	IPNS_FLASS
10	162	9.9	326	1	IPNS_LYSLA
11	143	8.7	333	1	IPNS_STRLP
12	140	8.6	348	1	FLS_PETHY
13	138	8.4	328	1	IPNS_NOCCLA
14	137.5	8.4	335	1	FLS_CITUN
15	137	8.4	321	1	IPNS_STRCT
16	134.5	8.2	349	1	FLS_SOLTU
17	133.5	8.2	331	1	IPNS_PENCH
18	127	7.8	334	1	FLS_EUSCR
19	122.5	7.5	331	1	IPNS_EMENI
20	121	7.4	356	1	LDOX_ARATH
21	118	7.2	365	1	FL3H_DIACA
22	113.5	6.9	337	1	FLS_MALDO
23	110	6.7	337	1	ISP7_SCHPO
24	109	6.7	336	1	FLS1_ARATH
25	108	6.6	336	1	FL3H_CALCH
26	106	6.5	338	1	IPNS_CEPAC
27	103.5	6.3	395	1	LDOX_MAIZE
28	103.5	6.3	430	1	LDOX_PETHY
29	101	6.2	357	1	LDOX_MALDO
30	99	6.0	358	1	FL3H_ARATH
31	96	5.9	357	1	FL3H_MATIN
32	96	5.9	364	1	FL3H_MALDO
33	95.5	5.8	291	1	FLS_MATIN

RESULT 1				
ID	CEFE_STRCL	STANDARD;	PRT;	311 AA.
AC	P18548;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOS)			
DE	(Expandase).			
GN	CEFE.			
OS	Streptomyces clavuligerus.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=1901;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;			
RX	MEDLINE=89123150; PubMed=2644235;			
RA	Kovacevic S., Weigel B.J., Tobin M.B., Ingolia T.D., Miller J.R.;			
RT	"Cloning, characterization, and expression in Escherichia coli of the			
RT	Streptomyces clavuligerus gene encoding deacetoxycephalosporin C			
RT	synthetase";			
RL	J. Bacteriol. 171:754-760(1989).			
CC	-I- FUNCTION: Catalyzes the step from penicillin N to deacetoxy-			
CC	cephalosporin C.			
CC	-I- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =			
CC	deacetoxycephalosporin C + succinate + CO(2) + H(2)O.			
CC	-I- COFACTOR: Iron and ascorbate.			
CC	-I- PATHWAY: Cephalosporin antibiotics biosynthesis.			
CC	-I- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF			
CC	OXIDOREDUCTASES. STRONG, TO CEFF.			

This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to license@isb-sib.ch).				

EMBL:	M32324;	AAA26715.1;	--	
DR	PIR;	T52312;	T52312.	
DR	PDB;	1DCS;	08-JUN-99.	
DR	PDB;	1ESH;	26-JUN-01.	
DR	PDB;	1ESI;	23-NOV-01.	
DR	PDB;	1HJF;	01-JUN-01.	
DR	PDB;	1HJG;	01-JUN-01.	
DR	PDB;	1RXP;	08-JUN-99.	
DR	PDB;	1RXG;	08-JUN-99.	
DR	InterPro;	IPR005123;	2OG-FeII_Oxy.	
DR	InterPro;	IPR002057;	Isopen_N_synth.	
DR	Pfam;	PF031171;	2OG-FeII_Oxy_I.	
DR	PROSITE;	PS00185;	IPNS_I;	FALSE_NEG.
DR	PROSITE;	PS00186;	IPNS_2;	1.
KW	Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C;			
KW	3D-structure.			

P33212 rhizobium f
P55712 rhizobium s
Q9RYV1 deinococcus
P24397 hyoscyamus
P51093 vitis vinif
Q07353 petunia hyb
P10967 lycopersico
P46550 caenorhabdi
P28038 hordeum vul
P41090 vitis vinif
P22248 azotobacter
Q8trk8 methanosarc

ALIGNMENTS

SQ SEQUENCE 311 AA; 34555 MW; 9C64E1FC37F524BC CRC64;
Query Match 99.5%; Score 1629; DB 1; Length 311;
Best Local Similarity 99.7%; Pred. No. 4.4e-143;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTVTFTSLAEIQQGLHQDEFFRCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFFHGSE 60
DB 1 MDTTVTFTSLAEIQQGLHQDEFFRCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFFHGSE 60
QY 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
DB 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
QY 121 QYFDROYTASRAVAREVLRTGTEPGGVAEFLDCEPLARFRYPQVPEHRSABEOPLM 180
DB 121 QYFDROYTASRAVAREVLRTGTEPGGVAEFLDCEPLARFRYPQVPEHRSABEOPLM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYNIRRTSKA 311
DB 301 NYNIRRTSKA 311

RESULT 2
CEFF_NOCLA STANDARD; PRT; 314 AA.
AC Q03047;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOCS)
DE (Expandase).
GN CEF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93173127; PubMed=8437592;
RA Coque J.J.R., Martin J.F., Liras P.;
RT "Characterization and expression in Streptomyces lividans of cefD and
RT cefE genes from Nocardia lactamdurans; the organization of the
RT cephamycin gene cluster differs from that in Streptomyces
RT clavuligerus";
RL Mol. Gen. Genet. 236:453-458(1993).
CC -1- FUNCTION: Catalyzes the step from penicillin N to deacetoxy-
CC cephalosporin C.
CC -1- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
CC deacetoxycephalosporin C + succinate + CO(2) + H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Cephalosporin antibiotics biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC
CC EMBL; Z13974; CAA78376.1; -
CC PIR; S30900; S30900.

DR InterPro; IPR005123; 2OG-FeII_Oxy.
DR InterPro; IPR002057; Isopen N synth.
DR Pfam; PF03171; 2OG-FeII_Oxy; 1.
DR PROSITE; PS00185; IPNS_1; FALSE NEG.
DR PROSITE; PS00186; IPNS_2; FALSE NEG.
KW Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
SQ SEQUENCE 314 AA; 34530 MW; D007AD3DEG7A1BDF CRC64;
Query Match 72.5%; Score 1186.5; DB 1; Length 314;
Best Local Similarity 69.8%; Pred. No. 3.6e-102; Indels 1; Gaps 1;
Matches 217; Conservative 41; Mismatches 52;
QY 2 DTTVTFTSLAEIQQGLHQDEFFRCLRDKGLFYLTDGCL-TDTLSAKDLVIDFFHGSE 60
DB 3 DATVTFTSLAEIQQGLHQDEFFRCLREKGVLYLKTGLPAEADHASGREIADVDFDHGTE 62
QY 61 AEKRAVTSVPVTRRGFTGLSESTAIQITNTGYSYDSCYMGTAADNLPSPGDFGRITW 120
DB 63 AEKKAAMTPIPTIRRGYAGLESESTAIQITNTGKYTDYSMSYMGTAADNLPSPAEFEKWE 122
QY 121 QYFDROYTASRAVAREVLRTGTEPGGVAEFLDCEPLARFRYPQVPEHRSABEOPLM 180
DB 123 DYPARMYRASQDVARQVLTSGAEPEVGMDFDCEPLRLRYFFPEVDRVAEOPLM 182
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGGAFDLPYRDPDAVLVFCGAIATLVGGQVKA 240
DB 183 APHYDLSIVTLIHQTPCANGFVSLQVEVDGSYVDIPAQFGLVFCGAVATLVADGAICA 242
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSFVPLARECGFDVSLDGETATFQDWIGG 300
DB 243 PKHVAAPGADKRVGSSRTSSVFFLRPNGDFFSVPRARECGFDVSIPTATFDDWIGG 302
QY 301 NYNIRRTSKA 311
DB 303 NYNIRKTA 313
RESULT 3
CEFF_NOCLA STANDARD; PRT; 310 AA.
AC F42219;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE Deacetylcephalosporin C synthetase (DACS) (Beta-lactam hydroxylase).
GN CEF.
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96287477; PubMed=8703431;
RA Coque J., Enguita F.J., Cardozo R.E., Martin J.F., Liras P.;
RT "Characterization of the cefE gene of Nocardia lactamdurans encoding
RT a 3'-methylcephem hydroxylase different from the 7-cepem
RT hydroxylase";
RL Appl. Microbiol. Biotechnol. 44:605-609(1996).
CC -1- FUNCTION: HYDROXYLATION OF DESACETOXYCEPHALOSPORIN C IN
CC 3'-POSITION TO FORM DEACETYLCEPHALOSPORIN C.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC
CC EMBL; Z13974; CAA78376.1; -
CC PIR; S30900; S30900.

```
DR EMBL; Z21687; CAA79803.1; -.
DR PIR; S40253;
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
DR Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
KW SEQUENCE 310 AA; 34364 MW; 9983EFAC1B41656 CRC64;
SQ
Query Match 59.0%; Score 965.5; DB 1; Length 310;
Best Local Similarity 61.2%; Pred. No. 9.4e-82;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;
QY 2 DTTVPFSLAELOOGLHODEFRCLDKGLFYLTDGGLTDTLTKSAKDLVIDFEHGSEA 61
DB 3 DKTVPFSLAELOOGLHODEFRCLDKGLFYLTDGGLTDTLTKSAKDLVIDFEHGSEA 61
QY 62 EKRAVTSPTVPTMRGFTGLSESTAOITNTGYSYDSCYSGMTADNLPSPGDFGRITWQ 121
DB 62 EKQAVTKVPTMRGYSALAEASTAQVTNTGTTDYSMSYSGMIGNLPSPKESFESVWTD 121
QY 122 YFDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFPYPPQVPEHRSAAEQPLRMA 181
DB 122 YFDSLRYAAQETARLVLTAAAGTVGDGDLTLLDCDPVLRLYPPEVPEHRAAEYEPERMA 181
QY 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTGGQVAP 241
DB 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTGGQVAP 241
QY 242 RHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWI 301
DB 242 NHVSPDASMLAGSDRTSSVFFLRPTDFTFSVPDARKYGLDVLDMKATPGDWIGTN 301
QY 302 YVNI 305
DB 302 YVTM 305
RESULT 4
CEFF_STRCL STANDARD; PRT; 318 AA.
AC P42270;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam hydroxylase).
GN CEFF.
OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1901;
RN [1]
RP MEDLINE=91100311; PubMed=1987130;
RA Kovacevic S., Miller J.R.;
RT "Cloning and sequencing of the beta-lactam hydroxylase gene (ceff)
RT from Streptomyces clavuligerus: Gene duplication may have led to
RT separate hydroxylase and expandase activities in the actinomycetes.";
RL J. Bacteriol. 173:398-400(1991).
CC -1- FUNCTION: HYDROXYLATION OF DESACETOXICEPHALOSPORIN C IN
CC 3' POSITION TO FORM DEACETYLCHEPHALOSPORIN C.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES. STRONG, TO CEFF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M63809; AAA26716.1; -.
DR
PIR; A39204; A39204.
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
DR Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
KW SEQUENCE 318 AA; 34584 MW; B17CC1CBC1E67178 CRC64;
SQ
Query Match 57.4%; Score 939.5; DB 1; Length 318;
Best Local Similarity 58.6%; Pred. No. 2.5e-79;
Matches 180; Conservative 43; Mismatches 81; Indels 3; Gaps 1;
QY 2 DTTVPFSLAELOOGLHODEFRCLDKGLFYLTDGGLTDTLTKSAKDLVIDFEHGSEA 61
DB 3 DTPVPIFNLALREGADQEKRECVTGMGVFLTGYGAGDKHRLATDTAMDFPANGTGA 62
QY 62 EKRAVTSPTVPTMRGFTGLSESTAOITNTGYSYDSCYSGMTADNLPSPGDFGRITWQ 121
DB 63 EKAAVTDVPTMRGYSALAEASTAQVTNTGTTDYSMSYSGMISGNVFPSPERFVWTE 122
QY 122 YFDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFPYPPQVPEHRSAAEQPL 178
DB 123 YFDKLYAAQETARLVLTASGGYDAEIVGSLDELLDADPVLRLRYPPEVPEHRSAAHEPR 182
QY 179 RNAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTGGQV 238
DB 183 RNAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPAVLVFCGAIATLVTGGQV 242
QY 239 KAPRHVAAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWI 298
DB 243 PAPRHVRSFGAGMREGSDRTSSVFFLRPTDFTFSVAKARSYGLAVLDMETATFGDWI 302
QY 299 GGNVNI 305
DB 303 GTNYVTM 309
RESULT 5
EXPA_CEPAC STANDARD; PRT; 332 AA.
AC P11935;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE Cephalosporin biosynthesis expandase/hydroxylase [Includes:
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOCS)
DE (Expandase); Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam
DE hydroxylase)].
GN CEFF.
OS Cephalosporium acremonium (Acremonium chrysogenum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; mitosporic Hypocreaceae;
OC Acremonium.
OX NCBI_TaxID=5044;
RN [1]
RP SEQUENCE FROM N.A.
RA Samson S.M., Dotzla J.E., Slisz M.L., Becker G.W., van Frank R.M.,
RA Veal L.E., Yeh W.K., Miller J.R., Queener S.W., Ingolia T.D.;
RT "Cloning and expression of the fungal expandase/hydroxylase gene
RT involved in cephalosporin biosynthesis.";
RL Biotechnology 5:1207-1214(1987).
CC -1- FUNCTION: DAOCS catalyzes the step from penicillin N to deacetoxy-
CC cephalosporin C, which is used as a substrate by DAOCS to form
CC deacetylcephalosporin C.
CC -1- CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
CC deacetoxycephalosporin C + succinate + CO(2) + H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Cephalosporin biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
CC
CC PIR; A29711; A29711.
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
```


complementation of the cephamycin pathway in Streptomyces
 RT Clavuligerus.";
 RL Antimicrob. Agents Chemother. 35:44-52(1991).
 CC -|- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
 CC from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to
 CC form the azetidinone and thiazolidine rings of isopenicillin.
 CC -|- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
 CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
 CC -|- COFACTOR: Iron and ascorbate.
 CC -|- PATHWAY: Biosynthesis of penicillin and cephalosporin.
 CC -|- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
 CC OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X54609; CAA38431.1; -;
 CC PIR; A61155; A61155.
 CC HSP; P05326; IAK0.
 CC InterPro; IPR005123; 2OG-Fell Oxy.
 CC InterPro; IPR002283; IPN synth.
 CC Pfam; PF031171; 2OG-Fell Oxy; 1.
 CC PRINTS; PR00682; IPNSYNTHASE.
 CC PROSITE; PS00185; IPNS_1; 1.
 CC Antibiatic biosynthesis; Oxidoreductase; Iron; Vitamin C.
 CC PROSITE; PS00186; IPNS_2; 1.
 CC METAL 212 212 IRON (BY SIMILARITY).
 CC METAL 214 214 IRON (BY SIMILARITY).
 CC METAL 268 268 IRON (BY SIMILARITY).
 CC SEQUENCE 329 AA; 37368 MW; 0CD96C8F7CF5A7EB CRC64;
 Query Match 10.8%; Score 176; DB 1; Length 329;
 Best Local Similarity 24.5%; Pred. No. 9,7e-09;
 Matches 73; Conservative 39; Mismatches 126; Indels 60; Gaps 12;
 Qy 5 VPTFSLAEQLQGLHQP-----BFRCLRDGKGLFYLTDCGLTDTLTELKAKDLVIDFHEG 58
 Db 10 VPTIDISPLSGDADKKRAQKVAQKACRESGFFYASHHGI---DVQLLDVVNEFHRTM 66
 Qy 59 SEAEK-----RAVTSVPVTRRGFTGLSESTAQITWTGSGYSYDSCYSGMTADNLF-- 111
 Db 67 TDBEKYDLAINAYNNKNNPRTRNGY-----YMAVKKKAVESWCYLNPSFSEDPQI 117
 Qy 112 -SG---DFGRW-----TQYFDRQYTSRAVAREVLRTGTEPDGGVEA--- 151
 Db 118 RSGTPMHEGNIWPEKHEQHFRCPECYRDVFSLSKVLKRGALALG-KPEDFFDASLS 176
 Qy 152 ----FLDCEPLLRFYFQVPEHRSABE-QPLRMAPHYDLSMVTLIQOTPCANGFVSLQAE 207
 Db 177 LADTSLAVTLIHYPYLDYDPVTKGDTKLSFEDHLDVSMITVLFQTEVQN----LQVE 232
 Qy 208 VGGFTDLPYRPDAVLVFCGAIATLVGGQVKAPRHVAAPRRDQIAGSSRTSSVPEL 265
 Db 233 TADGQDLPTSGENFLVNCGYMGYLTNDYFPAPNHRV-----KFVNAERLSLPFPL 284
 RESULT 8
 ID IPNS_STRCL STANDARD; PRT; 329 AA.
 AC P10621;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
 DE synthetase).
 GN PCBC.
 OS Streptomyces clavuligerus.

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1901;
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
 RX MEDLIN=88212175; PubMed=3130293;
 RA Leski B.K., Aharonowitz Y., Mevaresch M., Wolfe S., Vining L.C.,
 RT Westlake D.W.S., Jensen S.E.;
 RT "Cloning and nucleotide sequence determination of the isopenicillin N
 RT synthetase gene from Streptomyces clavuligerus";
 RL Gene 62:187-196(1988).
 CC -|- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
 CC from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to
 CC form the azetidinone and thiazolidine rings of isopenicillin.
 CC -|- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
 CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
 CC -|- COFACTOR: Iron and ascorbate.
 CC -|- PATHWAY: Biosynthesis of penicillin and cephalosporin.
 CC -|- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
 CC OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M19421; AAA26770.1; -;
 CC PIR; A01132; CAA00131.1; -;
 CC HSP; P05326; IBLZ.
 CC InterPro; IPR005123; 2OG-Fell Oxy.
 CC InterPro; IPR002283; IPN synth.
 CC Pfam; PF031171; 2OG-Fell Oxy; 1.
 CC PRINTS; PR00682; IPNSYNTHASE.
 CC PROSITE; PS00185; IPNS_1; 1.
 CC Antibiatic biosynthesis; Oxidoreductase; Iron; Vitamin C.
 CC PROSITE; PS00186; IPNS_2; 1.
 CC METAL 212 212 IRON (BY SIMILARITY).
 CC METAL 214 214 IRON (BY SIMILARITY).
 CC METAL 268 268 IRON (BY SIMILARITY).
 CC SEQUENCE 329 AA; 36958 MW; 71AA1CCE9514761C CRC64;
 Query Match 10.0%; Score 164; DB 1; Length 329;
 Best Local Similarity 23.8%; Pred. No. 1.2e-07;
 Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;
 Qy 5 VPTFSLAEQLQGLHQP-----BFRCLRDGKGLFYLTDCGLTDTLTELKAKDLVIDFHEG 58
 Db 10 VPTIDISPLSGDADKKRAQKVAQKACRESGFFYATNHGV---DVQQLDDVVNEFHGM 66
 Qy 59 SEAEK-----RAVTSVPVTRRGFTGLSESTAQITWTGSGYSYDSCYSGMT---A 106
 Db 67 TDOEKHDLAIHAYNPDPNPHVRNGYKAVGKAVESFCYLPDPFGEDHPM-IAAGTPMHE 125
 Qy 107 DNLFPDGD-----FGRITWQYFDRQYTSRAVAREVLRTGTEPDGGVEA-----FLDCE 156
 Db 126 VNLWDPDEERHPRPFCPEGYRQMLKSLVLMRGLALALG-RPEHFFDAALAEQDSLSV 184
 Qy 157 PLLRFRYFQVPEHRSABE-QPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGFTDL 215
 Db 185 SLIRYPYLEEYPPVKTPGQQLLSFEDHLDVSMITVLFQTEVQN-----LQVETVDMGRDI 240
 Qy 216 PYRPDAVLVFCGAIATLVGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNADFTSV 275
 Db 241 PTSENDFVNCGYMAHVTNDYFPAPNHRV-----KFVNAERLSLPFPLNGHEAVIE-- 293
 Qy 276 PLARECGFDVSLDGETATFQDWI 298

```

Db 294 PFVPE-GASEVNRNALSVDYL 315
RESULT 9
IPNS_FLASS STANDARD; PRT; 326 AA.
AC P16020;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
DE synthase).
DE PCBC.
OS Flavobacterium sp. (strain SC 12,154).
OC Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
OC Flavobacteriaceae; Flavobacterium.
OX NCBI_TaxID=241;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174998; PubMed=2308852;
RA Shiffman D., Cohen G., Aharonowitz Y., von Dohren H., Kleinkauf H.,
RA Nevarach M.;
RT "Nucleotide sequence of the isopenicillin N synthase gene (pcbC) of
RT the Gram-negative Flavobacterium sp. SC 12,154."
RL Nucleic Acids Res. 18:660-660(1990).
CC -!- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoacyl)-L-cysteiny-D-valine (ACV) to
CC form the acetidinone and thiazolidine rings of isopenicillin.
CC -!- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -!- COFACTOR: Iron and ascorbate.
CC -!- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -!- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X17355; CAA35233.1; -.
CC HSSP; P05326; 1BK0.
CC InterPro; IPR005123; 2OG-FeII Oxy.
CC InterPro; IPR002283; IPN synth.
CC Pfam; PF03171; 2OG-FeII Oxy; I.
CC PRINTS; PR00682; IPNSYNTHASE.
CC PROSITE; PS00185; IPNS_1; 1.
CC METAL 211 211 IRON (BY SIMILARITY).
CC METAL 265 265 IRON (BY SIMILARITY).
CC SEQUENCE 326 AA; 36465 MW; 256E274316395837 CRC64;
Query Match 9.9%; Score 162; DB 1; Length 326;
Best Local Similarity 22.9%; Pred. No. 1.9e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;
QY 24 RCLRDGLFYLTDCLGTLDELKSAKDLVDFEFGSEAEK-----RAVTSVPVPTMRGFT 78
Db 32 RACRSGGFFFAAHNGV---DLAALQKFTTDMHMAEKEWELAIRAYNPANPRNGY- 87
QY 79 GLESESTAQITNTGSDYSYCMYSGMTAD-----NLFPS-----GDFGRW 119
Db 88 -----YNAVEGKANESFCYLNPSFDADHATIKAGLPSHEVNIWPEARHPGMRFY 139
QY 120 TQVDFRQYATRAVAEVLNATGTEPDGGVEAF-----LDCEPLLRFYFQVPE 169
Db 140 EAYFSDVFDVAAILRGFAIALGRE-----ESFPERHFSMDDTLSAVSLIRYPFLENYP- 193

```

```

QY 170 HRSABEQPLRMAP-----HYDLSMTVLIQQTFCANGFVSLQAEVGAFTDLFPRD 220
Db 194 -----PLKGPDEKLSFEHHQDVLITLVLYQTAIPN-----LQVETAEGLDIPVSD 242
QY 221 AVLVFCAIATLVGGQVKAPRHVAAPRRQIAGSSRTSSVFF 264
Db 243 HFLVNGCTYMAHITNGYYPAPVHRV-----KYINAERLSIDPF 280
RESULT 10
IPNS_LYSLA STANDARD; PRT; 326 AA.
AC Q48739;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
DE synthase).
DE PCBC.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Lysobacter.
OX NCBI_TaxID=39596;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=YK90.
RX MEDLINE=96287475; PubMed=8703429;
RA Kimura H., Izawa M., Sumino Y.;
RT "Molecular analysis of the gene cluster involved in cephalosporin
RT biosynthesis from Lysobacter lactamgenus YK90."
RL Appl. Microbiol. Biotechnol. 44:589-596(1996).
CC -!- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
CC from delta-L-(alpha-aminoacyl)-L-cysteiny-D-valine (ACV) to
CC form the acetidinone and thiazolidine rings of isopenicillin.
CC -!- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
CC cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
CC -!- COFACTOR: Iron and ascorbate.
CC -!- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -!- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X56660; CAA39983.1; -.
CC FIR; S54099; S54099.
CC HSSP; P05326; 1BK0.
CC InterPro; IPR005123; 2OG-FeII Oxy.
CC InterPro; IPR002283; IPN synth.
CC InterPro; IPR002057; Isopen N synth.
CC Pfam; PF03171; 2OG-FeII Oxy; I.
CC PRINTS; PR00682; IPNSYNTHASE.
CC PROSITE; PS00185; IPNS_1; 1.
CC METAL 211 211 IRON (BY SIMILARITY).
CC METAL 265 265 IRON (BY SIMILARITY).
CC SEQUENCE 326 AA; 36564 MW; 25731F67173F8447 CRC64;
Query Match 9.9%; Score 162; DB 1; Length 326;
Best Local Similarity 22.9%; Pred. No. 1.9e-07;
Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;
QY 24 RCLRDGLFYLTDCLGTLDELKSAKDLVDFEFGSEAEK-----RAVTSVPVPTMRGFT 78
Db 32 RACRSGGFFFAAHNGV---DLAALQKFTTDMHMAEKEWELAIRAYNPANPRNGY- 87

```

QY 79 GLESESTAIQITNGSYSDYSCYMGATD-----NLFFS-----GDFGRW 119
 Db 88 -----YMAVEGKANESFCYLNPSFDADHATIKAGLPSHEVNIWDEARHPGMRFPY 139
 QY 120 TQYDFRQYTRASAVAREVLBATGTEDPGGVEAF-----LDCEPLRFRYPQVPE 169
 Db 140 EAYFSDVDVAAILRGAFAIALGRE-----ESFFERHFSMDTSLSAVSLIRYPFLENYP- 193
 QY 170 HRSAGEOPLRMAP-----HYDLSMTVLIQOTPCANGFVSLQAEVGGAFDLPVRPD 220
 Db 194 -----PLKLPDGEKLSFHHQDVSILVILVQTAIPN-----LQVETAGYLDIPVSDE 242
 QY 221 AVLVFCAIATLVGTGVKAPRHHVAAPRRDQIAGSSRTSSVFF 264
 Db 243 HFLVNCQTYMAHITNGYYPAPVHRV-----KVINAERLSIPFF 280

RESULT 11
 ID IPNS STRLP STANDARD; PRT; 333 AA.
 AC P12438;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N
 synthase).
 GN PCBC.
 OS Streptomyces lipmanii (Streptomyces alboniger).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=132472;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88314868; PubMed=3045077;
 RA Weigel B.J., Burgett S.G., Chen V.J., Skarud P.L., Prolik C.A.,
 Queener S.W., Ingolia T.D.;
 RT "Cloning and expression in Escherichia coli of isopenicillin N
 synthetase genes from Streptomyces lipmanii and Aspergillus
 nidulans";
 RT J. Bacteriol. 170:3817-3826(1988).
 RL CC
 CC -1- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms
 from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to
 form the azetidinone and thiazolidine rings of isopenicillin.
 CC -1- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-
 cysteinyl-D-valine + O(2) = isopenicillin N + 2 H(2)O.
 CC -1- COFACTOR: Iron and ascorbate.
 CC -1- PATHWAY: Biosynthesis of penicillin and cephalosporin.
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
 OXIDOREDUCTASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22081; AAA26771.1; -
 DR HSP; P05326; 1BK0.
 DR InterPro; IPR005123; 2OG-PeII Oxy.
 DR InterPro; IPR002283; IPN synth.
 DR InterPro; IPR002057; Isopen_N synth.
 DR Pfam; PF03171; 2OG-PeII_Oxy; 1.
 DR PRINTS; PR00682; IPNSYNTHASE.
 DR PROSITE; PS00185; IPNS 1; 1.
 DR PROSITE; PS00186; IPNS 2; 1.
 KW Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
 FT METAL 216 216 IRON (BY SIMILARITY).
 FT METAL 218 218 IRON (BY SIMILARITY).
 FT METAL 272 272 IRON (BY SIMILARITY).
 SQ SEQUENCE 333 AA; 38082 MW; D55385664EA2CA26 CRC64;

Query Match 8.7%; Score 143; DB 1; Length 333;
 Best Local Similarity 23.2%; Pred. No. 1.1e-05;
 Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;

QY 5 VTFSLAELOOGLHQP-----EPRCLRDKGLFVLTDCGLTDTLTKSAKOLVIDPFEH 57
 Db 10 VTFIDISPL-FGTDPDAKAVARQINEACRGSGFFYASHHGI---DVRRLQDVVNEFHRT 65
 QY 58 GSEAEK-----RAVTSPTVMERGFTGLESESTAIQITNGSYSDYSCY---SMGTADNL 109
 Db 66 MDOEKHDLAIHAYNNENSHVNGY-----YMARPKRTVESWCYLNPSFGEDHPM 116
 QY 110 FPSG-----DFGRITWQYDFRQYTRASRA-VAREVLRAVGT-----EPDG 147
 Db 117 IKAGTPMHEVNVMPDEERHPDERSFGQYRYREVFLSKVLLRGFALALGKPEPFENEV 176
 QY 148 GVEAFDCEPL-LRFRYPQVPE---HRSAGEOPLRMAPHYDLSMTVLIQOTPCANGFVS 203
 Db 177 TEEDTLSCRSLSMIRYPYLDYPPEAAIKTGPDTGLSFEDHLDVSMITVLVQTEVQN--- 232
 QY 204 LQAEVGGAFDLPYRDAVLVFCGAIATLVGTGVKAPRHHVAAPRRDQIAGSSRTSSVFF 263
 Db 233 LQVETVGGQSLTSGENFLINCGLYLTNDYFPAPNHRV-----KYVNAERLSLPP 286
 QY 264 FL 265
 Db 287 FL 288

RESULT 12
 ID FLS PETH STANDARD; PRT; 348 AA.
 AC Q07512;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flavanol synthase (EC 1.14.11.-) (FLS).
 GN FL.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.
 OX NCBI_TaxID=4102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Old Glory Blue; TISSUE=Petal;
 RX MEDLINE=94108485; PubMed=7904213;
 RA Holton T.A., Brugliera F., Tanaka Y.;
 RT "Cloning and expression of flavonol synthase from Petunia hybrida.";
 RL Plant J. 4:1003-1010(1993).
 CC -1- FUNCTION: CATALYZES THE FORMATION OF FLAVONOLS FROM
 DIHYDROFLAVONOLS. IT CAN ACT ON DIHYDROKAEEMPEROL TO PRODUCE
 KAEEMPEROL, ON DIHYDROQUERCETIN TO PRODUCE QUERCETIN AND ON
 DIHYDROMYRICETIN TO PRODUCE MYRICETIN.
 CC -1- COFACTOR: IRON AND ASCORBATE.
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR FLAVONOLS WHICH ARE
 FLAVONOIDS. FLAVONOLS ARE IMPORTANT CO-PIGMENTS IN FLOWER OR FRUIT
 COLOR AND ARE ALSO ESSENTIAL FOR POLLEN TUBE GROWTH.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGHEST LEVEL DURING THE FIRST
 STAGE OF FLOWER DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
 OXIDOREDUCTASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

```
DR EMBL; Z22543; CAA80264.1; -.
DR PUR; S33510;
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
KW Flavonoid biosynthesis; Oxidoreductase; Dioxxygenase; Iron; Vitamin C.
SQ SEQUENCE 348 AA; 39427 MW; B39E1E4381DE6379 CRC64;

Query Match      8.6%; Score 140; DB 1; Length 348;
Best Local Similarity 22.3%; Pred. No. 2.2e-05;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;

QY 27 RDKGLFVLTDCGLTDLTKSADLVDFRHSAGEAKRAVT-SPVPTMRGF-TGLESES 84
Db 77 KEGWIFQLNHGIPDENAIADLVKVGKEFFHVQEEKELIAKTPGSDIEGYTSLQKEV 136
QY 85 TAQITNTGSDYSYCMYSGMTADNLF-----PSGDFGRW-----TOYFDRQYT 128
Db 137 EKG-----KGMVDHFLFKIWPSSAVNRYWPKNPPSVREANEVEYGRMRE 181
QY 129 ASRAVAREVLRAATGPDGVEAFDCE--PLRFRYPQVQVPEHRSABEOPLRMAPHYDL 186
Db 182 VDRIRFKSLGLGLEGHEMIEAAGDEIVYLLKINYPPCPR-----PDALGVAHTDM 237
QY 187 SMYTLTQQTCCANGFVSQAEVCGAFTDLPYRPDAVLVFCGATATLVGTGQVKAPRHVA 246
Db 238 SYTIIIL-----VPEVQGLQVFKDGHVYDVXKIENALIVHGDQVEILSNKGKYSVTHRT 293
QY 247 APRDQIAGSSRTSSVFFLRPNAD 270
Db 294 VNK-----DKTRMSWPVLEPPSE 312

RESULT 13
IPNS_NOCLA
ID FLS_CITUN STANDARD; PRT; 328 AA.
AC Q27744;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isopenicillin N synthetase (EC 1.21.3.1) (IPNS) (Isopenicillin N synthase).
GN PCBC.
OS Nocardia lactamurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VAR LC 411;
RX MEDLINE=92065808; PubMed=1956290;
RA Coque J.J.R., Martin J.F., Calzada J.G., Liras P.;
RT "The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase, and pcbC of Nocardia lactamurans are clustered together in an organization different from the same genes in Arcanobium chrysogenum and Penicillium chrysogenum.";
RT in Arcanobium chrysogenum and Penicillium chrysogenum.";
RL Mol. Microbiol. 5:1125-1133(1991).
CC -1- FUNCTION: Removes, in the presence of oxygen, 4 hydrogen atoms from delta-L-(alpha-aminoadipyl)-L-cysteiny-D-valine (ACV) to form the acetidinone and thiazolidine rings of isopenicillin.
CC -1- CATALYTIC ACTIVITY: N-[(5S)-5-amino-5-carboxypentanoyl]-L-cysteiny-D-valine + O(2) -> isopenicillin N + 2 H(2)O.
CC -1- COFACTOR: Iron and ascorbate.
CC -1- PATHWAY: Biosynthesis of penicillin and cephalosporin.
CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF OXIDOREDUCTASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; X57310; CAA40562.1; -.
DR PIR; S15284; S15284.
DR HSSP; P05326; 1BK0.
DR InterPro; IPR005123; 2OG-FeII Oxy.
DR InterPro; IPR002283; IPN synth.
DR Pfam; PF03171; 2OG-FeII Oxy; 1.
DR PRINTS; PR00682; IPNSYNTHASE.
DR PROSITE; PS00185; IPNS 1; 1.
DR PROSITE; PS00186; IPNS 2; FALSE NEG.
KW Antibiotic biosynthesis; Oxidoreductase; Iron; Vitamin C.
FT METAL 210 210 IRON (BY SIMILARITY).
FT METAL 212 212 IRON (BY SIMILARITY).
FT METAL 266 266 IRON (BY SIMILARITY).
SQ SEQUENCE 328 AA; 37466 MW; F0DE8B0727AC3855 CRC64;

Query Match      8.4%; Score 138; DB 1; Length 328;
Best Local Similarity 22.6%; Pred. No. 3.1e-05;
Matches 67; Conservative 41; Mismatches 129; Indels 60; Gaps 12;

QY 5 VPTFSLAEQLQGLHQD-----EFRCLRDKGLFVLTDCGLTDLTKSADLVDFRHS 58
Db 8 VPTIDVSPFLGDDAQEKVRVGQEIKNACRGSGFFVAAHGV--DVQRLQDVVNEFHRTM 64
QY 59 SAEK-----RAVTFPVPTMRG-FGLESESTAQITNTGSDYSYCMY----- 101
Db 65 SPQEKYDLALHAYNNKNNSHVRNGYMAIEGKAVE-----SFCYLNPSFSDHPE 114
QY 102 -SMGT---ADNLPFGSD---EGRIWTOYFDRQYASRAVAREVLRAATGT-----EPDGG 148
Db 115 IKAGTFMHEVNSWPDEEKHPSFRPCEEYVWTWRLSKVLMRGFALALGKDERFFPELK 174
QY 149 VEAFLDCEPLRFRYPQVPEHRSABE-OPLRMAPHYDLSMYTLTQQTCCANGFVSQAE 207
Db 175 EADTLSSVSLIRYPVLEDDYPPVKTGPDGKLSFEDHFDVSMITLVYTQVQN---LQVE 230
QY 208 VCGAFTDLPYRPDAVLVFCGATATLVGTGQVKAPRHVAAPRRDQIAGSSRTSSVFF 264
Db 231 TVDGMWRDLTSDTDFLVNAGTYLGHLTNDYFPSPHLRV-----KFWNAERLSLPFF 281

RESULT 14
FLS_CITUN
ID FLS_CITUN STANDARD; PRT; 335 AA.
AC Q9ZMQ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flavonol synthase (EC 1.14.11.-) (FLS) (Citrus).
GN FLS.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Satsuma Mandarin;
RX PubMed=11903972;
RA Moriguchi T., Kita M., Ogawa K., Tomono Y., Endo T., Omura M.;
RT "Flavonol synthase gene expression during citrus fruit development.";
RL Physiol. Plantarum 114:251-258(2002).
CC -1- FUNCTION: CATALYZES THE FORMATION OF FLAVONOLS FROM DIHYDROFLAVONOLS. IT CAN ACT ON DIHYDROKAEEMFEROL TO PRODUCE KAEEMFEROL, ON DIHYDROQUERCETIN TO PRODUCE QUERCETIN AND ON DIHYDROMYRICETIN TO PRODUCE MYRICETIN (BY SIMILARITY).
CC -1- COFACTOR: IRON AND ASCORBATE (BY SIMILARITY).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR FLAVONOIDS WHICH ARE FLAVONOIDS. FLAVONOIDS ARE IMPORTANT CO-PIGMENTS IN FLOWER OR FRUIT COLOR AND ARE ALSO ESSENTIAL FOR POLLEN TUBE GROWTH.
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:34:58 ; Search time 41 Seconds
(without alignments)
1957.423 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637
Sequence: 1 MDTVTFTSLAEQLQGLHOD.....ATFQDWIGGNYNIRRTSKA 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_23:*
- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1369	83.6	311	2 Q93FD4	Q93fd4 streptomyc
2	937	57.2	319	2 Q48740	Q48740 lysobacter
3	917.5	56.0	332	3 Q9P4T5	Q9p4t5 cephalospor
4	852	52.0	313	2 Q48741	Q48741 lysobacter
5	168	10.3	366	16 Q987V6	Q987v6 rhizobium 1
6	161	9.8	329	2 Q9ANU0	Q9anu0 streptomyc
7	159	9.7	282	2 Q9FAC2	Q9fac2 streptomyc
8	156.5	9.6	403	5 Q9VCZ2	Q9vcz2 drosophila
9	156.5	9.6	403	5 Q8MZG6	Q8mzg6 drosophila
10	156	9.5	262	2 Q9FAC0	Q9fac0 streptomyc
11	154.5	9.4	366	10 Q942Y0	Q942y0 oryza sativ
12	154	9.4	262	2 Q9FAC1	Q9fac1 streptomyc
13	154	9.4	262	2 Q9FAB6	Q9fab6 streptomyc
14	152	9.3	262	2 Q9FAC3	Q9fac3 streptomyc
15	152	9.3	379	10 Q8VKL2	Q8vkl2 fagus sylv
16	151	9.2	262	2 Q9FAB8	Q9fab8 streptomyc

17	151	9.2	262	2 Q9FAB9	Q9fab9 streptomyc
18	151	9.2	331	2 Q9LCZ2	Q9lcz2 streptomyc
19	148.5	9.1	389	10 Q8RVF5	Q8rvf5 oryza sativ
20	146.5	8.9	362	10 Q80851	Q80851 arabidopsis
21	145.5	8.9	348	10 Q8SB89	Q8sb89 oryza sativ
22	144.5	8.8	376	10 Q9C955	Q9c955 arabidopsis
23	142.5	8.7	358	10 Q80850	Q80850 arabidopsis
24	140	8.6	262	2 Q9FAB7	Q9fab7 streptomyc
25	139.5	8.5	368	10 Q94LP7	Q94lp7 oryza sativ
26	136.5	8.3	380	10 Q39112	Q39112 arabidopsis
27	135.5	8.3	221	2 Q9F5K9	Q9f5k9 streptomyc
28	135.5	8.3	352	10 Q9LWJ3	Q9lwj3 oryza sativ
29	135	8.2	364	10 Q9LIF4	Q9lif4 arabidopsis
30	133	8.1	313	10 Q8H619	Q8h619 oryza sativ
31	131.5	8.0	355	16 Q8ZG11	Q8zgil yersinia pe
32	131	8.0	378	10 Q39111	Q39111 arabidopsis
33	130	7.9	309	16 Q8D029	Q8d029 yersinia pe
34	128.5	7.8	342	16 Q98NJ5	Q98nj5 rhizobium 1
35	128	7.8	346	10 Q8LP22	Q8lp22 nierenbergi
36	125	7.6	350	10 Q8H620	Q8h620 oryza sativ
37	125	7.6	382	10 Q04280	Q04280 phaseolus v
38	125	7.6	387	10 Q9LWJ4	Q9lwj4 oryza sativ
39	124.5	7.6	354	10 Q8SB87	Q8sb87 oryza sativ
40	124.5	7.6	370	10 Q04282	Q04282 phaseolus v
41	124	7.6	389	10 Q8L6W4	Q8l6w4 beta vulgar
42	124	7.6	392	10 Q9FS09	Q9fs09 malus domes
43	123.5	7.5	349	10 Q9LDV9	Q9ldv9 oryza sativ
44	123	7.5	339	10 Q40063	Q40063 hordeum vul
45	122.5	7.5	386	10 Q39541	Q39541 cucurbita m

ALIGNMENTS

RESULT 1

Q93FD4 ID Q93FD4 PRELIMINARY; PRT; 311 AA.
AC Q93FD4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE deacetoxycephalosporin C synthase.
OS Streptomyces jumoniensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29864;
RA Sim T.S., Sim J.;
RT "Cloning and purification of Streptomyces jumoniensis
RT deacetoxycephalosporin C synthase.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF317908; AAL09460.1; -
DR InterPro; IPR005123; 2OG-Fell_Oxy.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-Fell_Oxy; 1.
DR PROSITE; PS00186; IPNS_2; 1.
KW Forin.
SQ SEQUENCE 311 AA; 34242 MW; B821ED796A4F5C1A CRC64;

Query Match 83.6%; Score 1369; DB 2; Length 311;
Best Local Similarity 81.4%; Pred. No. 1.4e-119;
Matches 253; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

QY	1	MDTVTFTSLAEQLQGLHODFRRCRLDRDGLFYLTDCGLTDTLTKSAKDLVDFFFHSGE	60
Db	1	MDTVTFTSLAEQLQGLHODEFRSCLAEGLFYLTDSLGLSDADQKSAKDAVIDFFEHGTE	60
QY	61	AEKRAVTSVPVPMRGFTGLESESTAIQTNTGSDYSNCYSNGTADNLPSPGDFGRWT	120
Db	61	EKRAATSTIPTIRRGFTGLESESTAIQTNTGSDYSNCYSNGLADNVFPFGDFERVWT	120

```
QY 121 QYFDROYTASRAVAREVLATGTGPGVEAFDCEPLLRFRYFPQVPEHRSABEQPLRM 180
DB 121 HYFGMYDVTOEVARVQLTKTGTGPEVGVESFVDCPEPLLRFRYFPVPEHRSABEQPLRM 180
QY 181 APHYDLSWVTLIIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIAITLVGGQVKA 240
DB 181 APHYDLSWVTLIIQOTPCNGFVSLQAEIGAFVLDLPAKFDPAVLVFCGAIAITLVGGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 300
QY 301 NYVNIIRTSKA 311
DB 301 NYVNIIRTSKA 311

RESULT 2
Q48740
ID Q48740 PRELIMINARY; PRT; 319 AA.
AC Q48740;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Deacetylcephalosporin C synthetase.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Lysobacter.
OX NCBI_TaxID=39596;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YK90.
RX MEDLINE=96287475; PubMed=8703429;
RA Kimura H., Izawa M., Sumino Y.;
RT "Molecular analysis of the gene cluster involved in cephalosporin
RL biosynthesis from Lysobacter lactamgenus YK90.";
RL Appl. Microbiol. Biotechnol. 44:589-596(1996).
DR EMBL: X56660; CAA39984.1; -.
DR InterPro: IPR005123; 2OG-Fell_Oxy.
DR Pfam: PF03171; 2OG-Fell_Oxy; 1.
KW Porin.
SQ SEQUENCE 319 AA; 35559 MW; 51040CD201B7C272 CRC64;

Query Match 57.2%; Score 937; DB 2; Length 319;
Best Local Similarity 58.0%; Pred. No. 3.5e-79;
Matches 177; Conservative 45; Mismatches 83; Indels 0; Gaps 0;

QY 2 DTTVPTFSLAELOQGLHQDEFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEGHSEA 61
DB 3 DSGIQIFDLDEHGVRLDSFRKSLPERGVYFYREDSDSIKTEHAKAMDAVMDLFFENGSAE 62
QY 62 EKRAVTSVPVPMRGFTGLESESTAQITNTGSDYSYCMYSGMTADNLPSPGDFGRITQ 121
DB 63 QKNALRNLTENVRRGFSDLAEASTARITKGGEVTDYSMVYSIGLTDNLPSPAPFAEITWG 122
QY 122 YFDROYTASRAVAREVLATGTGPGVEAFDCEPLLRFRYFPQVPEHRSABEQPLRM 181
DB 123 YFDRFYARTQDIARSVLRASDAGLSDVDFDLPLRFRFPVPEHRSABEQPLRM 182
QY 182 PHYDLSWVTLIIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIAITLVGGQVKA 241
DB 183 PHYDLSWVTLIIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIAITLVGGQVKA 242
QY 242 RHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 301
DB 243 RHQVAPSPMQVRGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 302
QY 302 YVNIIR 306
DB 303 YVNIIR 307

RESULT 3
```

```
Q9P4T5
ID Q9P4T5 PRELIMINARY; PRT; 332 AA.
AC Q9P4T5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Deacetylcephalosporin C synthetase/hydroxylase.
GN CEFEF.
OS Cephalosporium acremonium (Acremonium chrysogenum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acremonium.
OX NCBI_TaxID=5044;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3/2;
RA Jekosch K., Nosek J., Kueck U.;
RT "Analysis of cefer gene expression in two cephalosporin C producing
RL strains of Acremonium chrysogenum.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ404737; CAB96750.1; -.
DR InterPro: IPR005123; 2OG-Fell_Oxy.
DR InterPro: IPR002057; Isopen_N_synth.
DR Pfam: PF03171; 2OG-Fell_Oxy; 1.
DR PROSITE: PS00185; IPNS_1; 1.
DR PROSITE: PS00186; IPNS_2; 1.
KW Porin.
SQ SEQUENCE 332 AA; 36538 MW; E3DECESE9C28879 CRC64;

Query Match 56.0%; Score 917.5; DB 3; Length 332;
Best Local Similarity 56.4%; Pred. No. 2.4e-77;
Matches 176; Conservative 43; Mismatches 92; Indels 1; Gaps 1;

QY 1 MDTTPTFSLAELOQGLHQDEFRCLRDKGLFYLTDCGLTDTLKSADLVDPFEGHSEA 60
DB 1 MTSKVPVFLRDLKSGKVLTELAETTKGIFYLTESGLVDDHDSARETCVDFPKNGSE 60
QY 61 AEKRAVTSVPVPMRGFTGLESESTAQITNTGSDYSYCMYSGMTADNLPSPGDFGRITQ 120
DB 61 EKRATVTLADRNARRGFSNALEWSTAVVTETGYSYCMYSGIGGNLFPNRPEDVWQ 120
QY 121 QYFDROYTASRAVAREVLATGTGPGVEAFDCEPLLRFRYFPQVPEHRSABEQPLRM 179
DB 121 DYFDRMYGAADVARAVLNSVGAFLAGEDIDDFEVDCLLRLYRFPVPEHRSABEQPLRM 180
QY 180 MAPHYDLSWVTLIIQOTPCANGFVSLQAEVGAFTDLPYRDAVLVFCGAIAITLVGGQVKA 239
DB 181 MGPHYDLSITLIVHOTACANGFVSLQAEVGAFTDLPYRDAVLVFCGAIAITLVGGQVKA 240
QY 240 APRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 299
DB 241 APKRVKSPGRDQVRGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFQDWIG 300
QY 300 GNYVNIIRTSKA 311
DB 301 GNYVNIIRTSKA 312

RESULT 4
Q48741
ID Q48741 PRELIMINARY; PRT; 313 AA.
AC Q48741;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Deacetylcephalosporin C synthetase.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Lysobacter.
OX NCBI_TaxID=39596;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YK90;
RX MEDLINE=96287475; PubMed=8703429;
```

```

RESULT 5
Q987V6 PRELIMINARY; PRT; 365 AA.
AC Q987V6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DE Hypothetical protein mlr6892.
GN MLR6892.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti ";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003010; BAB53094.1; -.
DR InterPro; IPR005123; 2OG-FcII Oxy.
DR InterPro; IPR002283; IPR synth.
DR Pfam; PF03171; 2OG-FcII Oxy; 1.
DR PRINTS; PR006682; IPRNSYNTHASE.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 365 AA; 40188 MW; 67EDD760AFFDC0C9 CRC64;
Query Match 10.3%; Score 168; DB 16; Length 366;

```

<div></div>						
Best Local Similarity 24.7%; Pred.No. 3.2e-07;						
Matches	85; Conservative	41; Mismatches	158; Indels	60; Gaps	14;	
QY	1	MDTTFSTFSLAEIQQGLHQ-----DFRRCLRDKGLFYLTDCGLTDTELKSADLVIDFF	55			
Dd	28	MPRIVPVLDSLRLAQAGSERRTFLADLSASRDIGPFYLGHGISMWIEISVLTASFQQFF	87			
QY	56	EHGSEAKKAIVTSPTVMRGFTGLESESTAQTNTGSYSYMCMYSGMGTADNLFPSPGDP	115			
Dd	88	AL--PEADKLAIEMVKSSQFRGYTRAGELTK-----GREDMRELQDIGVERQAIAQPG	140			
QY	116	GRIWTQ-YPDROYTAS-----RAVAREVLRTATGEPPDGVEAFLDCEPILL	159			
Dd	141	TPAWTLQGNQPAAUPLDLKPALLAOWSKVTAVAIRLLKFAQSLODQPEDAF---DPI-	196			
QY	160	RFRYPFPQVEHR-SABEQPLR-----MAPHYDLSMWTLTIQTCPANGFVSIOAEVGG	210			
Dd	197	----YSISPENHMKVIYVRPGRTDTGGDGVGAHKDGGFLTLLQ-----DDNKGLQVDYG	248			
QY	211	AFTDLPRPDALVFCAITATLTVGGOVKAPRHVVAAARRDDQTAGSRTSSVFLRPNAD	270			
Dd	249	SWDVDPIPCTLVVNICELLEIASNGYLRTAVRVVTPP-----AGVERISVPPFFFSARLD	304			
QY	271	FT-----FSVPLEARECGFDVSLDETATTODWIGGNNVVNRITS	309			
Dd	305	ATIPLLGLSELAAARGPAS-DPDNPFLFRD-VGTNVLKSRLLRS	346			
RESULT 6						
ID	Q9ANU0	PRELIMINARY;	PRT;	329 AA.		
AC	Q9ANU0;					
DT	01-JUN-2001	(TrEMBLrel. 17, Created)				
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)				
DE	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)				
Dt	IsoPenicillin N synthase.					
GN	PCBC.					
OS	Streptomyces fimbriatus.					
OC	Bacteria; Actinobacteria; Actinomycetaceae; Streptomycetales;					
OX	Streptomycineae; Streptomycetaceae; Streptomycetes.					
NCBI_TaxID=68197;						
[1]						
SEQUENCE FROM N.A.						
RC	STRAIN=JCM 4910;					
RA	Sim T.S., YO C.Y., Wong E.;					
RT	"Genome walking and PCR-based cloning of isopenicillin N synthase gene from Streptomyces fimbriatus";					
RL	Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.					
DR	EMBL; AF320779; AAK11177.1; -					
HSP;	P05326; IBLZ.					
DR	InterPro; IPR005123; ZOG-Fell_Oxy.					
DR	InterPro; IPR002283; IPN synth.					
DR	InterPro; IPR002057; IsoPen_N_synth.					
DR	Ffam; PF03171; ZOG-Fell_Oxy; I.					
DR	PRINTS; PR00682; IPNSYNTHASE.					
DR	PROSITE; PS00185; IPNS 1; 1.					
DR	PROSITE; PS00186; IPNS 2; 1.					
SQ	SEQUENCE 329 AA; 37232 MW; 4BC3CA09CC556A9 CRC64;					
Query Match 9.8%; Score 161; DB 2; Length 329;						
Best Local Similarity 23.4%; Pred.No. 1.2e-06;						
Matches	70; Conservative	42; Mismatches	125; Indels	62; Gaps	13;	
QY	5	VPTFSLAEL-----QQGLH-QDEFRRCLRDKGLFYLTDCGLTDTELKSADLVIDFFEHG	58			
Dd	10	VPTDISPLFGDDPDPAKTHVAQQINKACRSGGFYASHHI-----DVQOLQDVNFEGHTM	66			
QY	59	SEAEEK-----RAVTSVPVTTMRRG-FTGLESESTAQTINTGSYSYSCMYC-----SMGTADNL	109			
Dd	67	TDEEKDYDLAINAYNSANPRVNGYMAVECKAVE-----SWCYLNPSFGEDHPM	116			
QY	110	FPSGD-----FGRITWQYFDRQYTASRAVARVLRATGTEPDGGEA--	151			

```

Db      117  IIRSGTGMHEVNIWPKDEKHERFRPFCEQYYRDMFQSLKTLMRGFALALG-KPEDFFDANL 175
QY      152  ----FLDCPLLRFRFPQVBEHRSABE-QPLEMAHYDLSWVTLIQOTPCANGFVSLQA 206
Db      176  PEDDTLSAVSLIRYPHKLKAYPPVKTPGDTGKUSFEDHLDVSVITLVLFQTEVQN----LQV 231
QY      207  EVGGAFTDLPYRPDAVLVFCGAIATLVTGQVKAPRRHVAAPRRDQIAGSSRTSSVFEL 265
Db      232  ETVNGHQDLFTSGDDFLVNGCTYMGVLTNDYPPAPNHRV-----KFINARLSLPPFL 284

RESULT 7
Q9FAC2  PRELIMINARY; PRT; 262 AA.
AC      Q9FAC2;
DT      01-MAR-2001 (TREMBlrel. 16, Created)
DT      01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Isopenicillin N synthase (IPNS) (Fragment).
GN      PCBC.
OS      Streptomyces panayensis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycineae; Streptomycetaceae; Streptomycetes.
OX      NCBI_TaxID=132470;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=JCM 5042.
RA      Palaniappan N., Seki T.;
RT      "Phylogenetic relationship between the Isopenicillin N synthase (IPNS)
RL      and 16S rDNA genes.";
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB045851; BAB13299.1; -
DR      HSSP; P05326; 1BK0.
DR      InterPro; IPR005123; 2OG-FeII Oxy.
DR      InterPro; IPR002283; IPN synth.
DR      InterPro; IPR002057; Isopen N synth.
DR      Pfam; PF03171; 2OG-FeII Oxy; 1.
DR      PRINTS; PR00682; IPNSYNTHASE.
DR      PROSITE; PS00185; IPNS 1; 1.
DR      PROSITE; PS00186; IPNS 2; 1.
FT      NON TER 1
FT      NON TER 262
SQ      SEQUENCE 262 AA; 30074 MW; D9A5D0BAB942D776 CRC64;

Query Match 9.7%; Score 159; DB 2; Length 262;
Best Local Similarity 25.4%; Pred. No. 1.4e-06;
Matches 63; Conservative 38; Mismatches 115; Indels 32; Gaps 10;

QY      21  EFRCLRDKGLFYLTDCGLTDTBLKSAKDLIVDFEHSSEAEK-----RAVTSVPVPTMR 75
Db      17  EINKACRGSGFFVASHHGI---DVQLLDVVNEFHRTMTDEEKYDLAINAYKNKNPHVN 73
QY      76  GF-----TGLESESTAQITNGSYSDYSCVSMGT---ADNLFPS-----GDFGRITWQYFD 124
Db      74  GYTMVKGKKAVERSVLNRPSFSEDFHMIRS-GTPMHEVNIWPKDEKHERFRPFCEQYYR 132
QY      125  RQYTSARAVAREVLRTGTPEDCGVGA-----FLDCEPLLRYPFPQVPEHRSABE-QP 177
Db      133  DMFQLSKALMRGFPALALG-KPEDFFRAPLSQDTLSAVSMIRPYLLEDPPVKVTGPDGTYK 191
QY      178  LRMAPHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVTTGQ 237
Db      192  LSFEDHLDVSMITVLVLFQTEVQN----LQVETVDGWDLPSTGSENFVLCNGTYMGVLTNDY 247
QY      238  VKAPRRHV 245
Db      248  FPAPNHRV 255

RESULT 8
Q9VCZ2  PRELIMINARY; PRT; 403 AA.
ID      Q9VCZ2
AC      Q9VCZ2;

```

```

DT      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      CG5346 protein.
GN      CG5346.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Berkelley;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA      Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harlow N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Swirekas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195(2000).
RN      [2]
RC      SEQUENCE FROM N.A.
RA      Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA      Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA      Banson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
RA      Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA      Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA      Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA      Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA      Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA      McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA      Pacleb J., Paragaa V., Park S., Patel S., Pfeiffer B.,
RA      Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA      Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA      Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT      "Sequencing of Drosophila melanogaster genome.";
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RC      SEQUENCE FROM N.A.
RP      Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA      Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,

```

[illegible]

DR	ENBL; AY102699; AAM27528.1.; -.
DR	FlyBase; FBgn038981; CGS346.
DR	InterPro; IPR005123; ZO-Feil_Oxy.
DR	Pfam; PF03171; ZO-Feil_Oxy; 1.
SQ	SEQUENCE 403 AA; 45461 MW; DE2CBC0FD443A07C CRC64;
Query Match	9.6%; Score 156.5; DB 5; Length 403;
Best Local Similarity	20.2%; Pred. No. 4.3e-06;
Matches 72; Conservative 51; Mismatches 127; Indels 107; Gaps 12;	
QY	3 TTVPTESLAELOQLHQ-----DEFRCRLDKGLFYLTDCLGTDTLTKSAKDLV 51 : : : : : : : : :
DB	18 SVVPIDLAHC--GIEPVPKVSVNRVGHLKKALSEKGHALLVNHGISEKLKTANDHL 75 : : : : : : : : :
QY	52 IDPFHEGSEAERAVTSPVTMRGRFTGLESESTAQTN-----TGSYSDYSMCYSNGT 105 : : : : : : : :
DB	76 DDF-----VNLP-PDIRQHVIADGDGHGVSRQQORFGDCKSPELRHAFNI 123 : : : : : : : : :
QY	106 --ADNL-----FPSGDFRITWTFDRQYTASRAVAREVLRATCTEPGGVEAPL- 153 : : : : : : : :
DB	124 LNAQNLFEEPLPG-----FADHI STLATDFKALASFILQALVASLDIPHTFFLEKISHM 177 : : : : : : : : :
QY	154 -----DCEPLLRFYPQV-----PEHRSAEEOPL 178 : : : : : : : : :
DB	178 LSGDHNMSSLRMLYPIVDDEBPQGNDVKRGCRQYSQRCLSNQPDPFRPHNPREDDDL 237 : : : : : : : : :
QY	179 -----RNAPHYDLSMTVLTIQOTPCANGFVSQAEGVGAAFTDLPY 217 : : : : : : : : :
DB	238 NEVDGPNGLQFEHLGNLGNVIGCPHVVDYGFILLSD--SEGLEVLPSCSKWNVRGH 295 : : : : : : : : :
QY	218 RPDALVFCAIATLVTTGGVKAPRRHHVAAPRDQIAGSSRTSVFLRPNADTPFS 274 : : : : : : : : : : : :
DB	296 LPGSILNVCGEILNIWTQGRYPALQHRVIIPEQETIRARGRHSTAFFCHPDNITTIS 352 : : : : : : : : : : : :
RESULT 10	
Q9FAC0	PRELIMINARY; PRT; 262 AA.
ID Q9FAC0	
AC Q9FAC0	
DT 01-MAR-2001 (T-EMBLrel. 16, Created)	
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)	
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)	
DE Isopenicillin N synthase (IPNS) (Fragment).	
DN PCBC.	
GN Streptomyces heteromorphus.	
OS Bacteria; Actinobacterias; Actinomycetales;	
OC Streptomycinae; Streptomycetaceae; Streptomyces.	
RX NCBI_TaxID=132471;	
ON [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=JCM 5031;	
RA Palaniappan N., Seki T.;	
RT "Phylogenetic relationship between the Isopenicillin N synthase (IPNS)	
RL and 16S rDNA genes.";	
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.	
ENBL; AB045853; BAB13301.1; -.	
HSP; P05326; IBKO.	
DR InterPro; IPR005123; ZO-Feil_Oxy.	
DR InterPro; IPR002283; IPN_synth.	
DR InterPro; IPR002057; Isopen_N synth.	
DR Pfam; PF03171; ZO-Feil_Oxy; 1.	
DR PRINTS; PR00682; IPNSYNTHASE.	
DR PROSITE; PS00185; IPNS_1; 1.	
DR PROSITE; PS00186; IPNS_2; 1.	
FT NON TER 1 1	
FT NON TER 262 262	
SQ SEQUENCE 262 AA; 30122 MW; 3AFED20F3DB41B69 CRC64;	
Query Match	9.5%; Score 156; DB 2; Length 262;
Best Local Similarity	25.0%; Pred. No. 2.6e-06;
Matches 62; Conservative 38; Mismatches 116; Indels 32; Gaps 10;	
OY 21 EFRCRLDKGLFYLTDCGLTDTLKAKDLVIDPFHEGSEAEEK-----RAVTSVPVTMR 75	

DR	PROSITE; PS00185; IPNS 1; 1.
DR	PROSITE; PS00186; IPNS 2; 1.
FT	NON_TER 1_262
FT	NON_TER 262
SQ	SEQUENCE 262 AA; 30107 MW; D72BF9737CC8EAB CRC64;
 Query Match Best Local Similarity 9.3%; Score 152; DB 2; Length 262; Matches 62; Conservative 39; Mismatches 115; Indels 32; Gaps 10;	
Qy	21 EFRCLRDKGLFYLDCGLTDTLKSADLVDPFEHSGAEK-----RAVTSVPVTMR 75
Db	17 EINKACRGSGFFFYASHHG1---DVQLQDVVNEFHRTMTDEEKHELAIHAYNQANPRVN 73
Qy	76 GF-----TGLESESTAQTINTGSYSDSMCYSMGT---ADNLFPSSGD----FGRIWTOYFD 124
Db	74 GYYMAVGKKAVESWCYNLPAFGEDHPMIRS-CTPLHEVINWDEKRHERFRPFCQYR 132
Qy	125 RQYTASRAVARVLRTATGTEPDGGVAFL---DCEPLLRFYPQPVEHRSBEQP----- 177
Db	133 DMFGLSKTLMRGFALALG-KPEDFFDAHLPEADTLSAVSLIRYPRLEDYPPVKTPGDKTK 191
Qy	178 LRMAPHYDLISWVTLIQOTPCANGFSVLSQAEVGAFTDLPRPDAVLVFCALATLVTTGGQ 237
Db	192 LSFQDHLDSMITVLFQTEVQN----LQVETVDGWQDLPTSGDFLVNCGTFMGYLNDY 247
Qy	238 VKAPRRHV 245
Db	248 FPAPNRV 255
 RESULT 15	
ID	QBVXL2 PRELIMINARY; PRT; 379 AA.
AC	QBVLX2
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Gibberellin 20-oxidase 1.
GN	GA20OX1.
OC	Fagus sylvatica (Beechnut).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid 1; Fagales; Fagaceae; Fagus.
ON	NCBI_TaxID=28930;
RX	[1]
RP	SEQUENCE FROM N.A.
RT	Calvo A., Lorenzo O., Nicolas C., Rodriguez D., Nicolas G.;
RT	"Molecular cloning and transcript analysis of a functional GA20
RT	oxidase related to the release from dormancy in Fagus sylvatica
RT	seeds.";
RL	Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AJ420192; CAD21846.1; -.
DR	InterPro; IPR005123; ZOG-FcII_Oxy.
DR	InterPro; IPR002283; IPN synth.
DR	Pfam; PF031171; ZOG-FcII_Oxy; 1.
DR	PRINTS; PR00682; IPNSYNTHASE.
SQ	SEQUENCE 379 AA; 43056 MW; E3E51A2FC94EFBB3 CRC64;
 Query Match Best Local Similarity 24.3%; Score 152; DB 10; Length 379; Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;	
Qy	30 GLFYLDCGLTDTLKSADLVDPFEHSGAEKRAVTSVPVTMR-----GFTG- 79
Db	92 GFFLVNVHGDDKLIAHAHQYIDYFFELPMASKORA-----QRKVGEHCGVASSFTGR 144
Qy	80 -----LSESESTAQTINTGSYSDSMCYSMGTADNLFPSSGDFGRIWTOYFDROYTA 129
Db	145 FSKLPWKETLSRPSAQPDSSINIVQDY-LCNTWG--EDFKP---FGKVTQDYCDAMSTL 198
Qy	130 SRARAREVLRTATEPDGGVAEFLDCEPLLRFYFP--QVPEHRSAEQPLRMAPHYDLS 187

Db	199	SLGIMELLGMSLGVSGHYREFEENESINRLNYPYCQKP-----DLTGTGPHCDPT	252
Qy	188	MVTLIQOTPCANGFVSLQAEVGG--AFTDLPYRP-----DAVLVFCGAIATLVTGGQVKA	240
Db	253	SLTILHQD-----QVGGLOVFEDEWRSITPNFNAFVVNIGDTFMALSNGRYKS	301
Qy	241	PRHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVP	276
Db	302	CLHRAVVNSK-----TPRKSLAFFLCPCKNDKVVSPP	332

Search completed: February 3, 2004, 17:38:34
Job time : 44 secs


```
Db 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFDWIGG 300
QY 301 NYVNIIRTSKA 311
Db 301 NYVNIIRTSKA 311

RESULT 2
US-08-592-411-15
; Sequence 15, Application US/08592411
; Patent No. 5726032
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for the Efficient Production of
; TITLE OF INVENTION: 7-ADCA via 2-(Carboxyethylthio)acetyl-7-ADCA and
; TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,411
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-592-411-15

Query Match 73.4%; Score 1201; DB 1; Length 313;
Best Local Similarity 70.3%; Pred. No. 3.3e-128;
Matches 218; Conservative 41; Mismatches 51; Indels 0; Gaps 0;

QY 2 DTTVPTFSLAELOOGLHODEFRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFPHGSEA 61
Db 3 DATVPTFDLAELEGLHQEFREHCLREKGVFLYLGTLGTPAEADHASGREIAVDFFDHGTE 62
QY 62 EKRAVTSVPVPTMRGFTGLSESTAQITNTGSDYSMYSGMTADNLPSPGDFGRITWQ 121
Db 63 EKXAVMTPIIRRGYAGLESESTAQITNTGKYDYSMSYSGMTADNLPSPAEFEKAWED 122
QY 122 YFDRQYTASRAVAREVLRAATGTPDGGVEAFDCEPLLRFRYPQVPEHRSABEQPLRMA 181
Db 123 YFARMYRASQDVARQVLTSGAEPEVGMDFDCEPLLRFRYPQVPEHRSABEQPLRMA 182
QY 182 PHYDLSMVTLLIQOTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGTGGVKAP 241
Db 183 PHYDLSIVTLIHQTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAVATLVADGAIKAP 242
QY 242 RHVVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFDWIGG 301
Db 243 KHVVAAPGADKRVGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFDWIGG 302
QY 302 YVNIIRTSKA 311
Db 303 YINIRKTA 312

RESULT 3
US-08-592-411-17
; Sequence 17, Application US/08592411
; Patent No. 5726032
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for the Efficient Production of
; TITLE OF INVENTION: 7-ADCA via 2-(Carboxyethylthio)acetyl-7-ADCA and
; TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA
; NUMBER OF SEQUENCES: 17
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,411
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-592-411-17

Query Match 72.5%; Score 1186.5; DB 1; Length 314;
Best Local Similarity 69.8%; Pred. No. 1.5e-126;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOOGLHODEFRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFPHGSEA 60
Db 3 DATVPTFDLAELEGLHQEFREHCLREKGVFLYLGTLGTPAEADHASGREIAVDFFDHGTE 62
QY 61 EKRAVTSVPVPTMRGFTGLSESTAQITNTGSDYSMYSGMTADNLPSPGDFGRITW 120
Db 63 EKXAVMTPIIRRGYAGLESESTAQITNTGKYDYSMSYSGMTADNLPSPAEFEKAW 122
QY 121 QYFDRQYTASRAVAREVLRAATGTPDGGVEAFDCEPLLRFRYPQVPEHRSABEQPLRM 180
Db 123 DYFARMYRASQDVARQVLTSGAEPEVGMDFDCEPLLRFRYPQVPEHRSABEQPLRM 182
QY 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAIATLVGTGGVKA 240
Db 183 APHYDLSIVTLIHQTPCANGFVSLQAEVGGATDLPYRPDAVLVFCGAVATLVADGAIK 242
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFDWIGG 300
Db 243 KHVVAAPGADKRVGSSRTSSVFFLRPNADFTFSPVLARECGFDVSLDGETATFDWIGG 302
QY 301 NYVNIIRTSKA 311
Db 303 YINIRKTA 312

RESULT 4
PCT-US95-04801-6
; Sequence 6, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan P.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaurena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
```

```
/
/
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Wallen III, John W.
/ REGISTRATION NUMBER: 35,403
/ REFERENCE/DOCKET NUMBER: 19179
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 594-3905
/ TELEFAX: (908) 594-4720
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 310 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ PCT-US95-04801-6

Query Match 59.0%; Score 965.5; DB 5; Length 310;
Best Local Similarity 61.2%; Pred. No. 2.1e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPFSLAELQGLHODEFRCLRDKGLFYLTDCGLTDTLTELKSAKDLVIDFEHSE 61
DB 3 DKTVPVFSMAELRDSGRQDEFWAR-RGVFYLTGYGATERDHRVATDTAMDFPAQTAE 61

QY 62 EKRAVTSVPVPMRRGFTGLESESTAOITNTGSDYSCMYSMGTADNLPSPSGDFGRIWTQ 121
DB 62 EKQAVTKVPTMRGVSALREASTAQVTNTGTTIDYMSYSGMIGNLPFSKESFESVWTD 121

QY 122 YFDRQYTTASRAVAREVLRTATGTPDGGVEAFDCEPLLRFPYQVPEHRSAAEQPLRMA 181
DB 122 YFDSLRYRAAQETARLVLTAAAGTYDGEDLDLTLLDCDPVLRRLRYFPEVPEHRAAEYEPERMA 181

QY 182 PHYDLSMVTLIQOTPCANGFVSLQAEVGGAFDLPYRPPDAVLVFCGAIATLVTGGQKAP 241
DB 182 PHYDLSIITPIHTPCANGFVSLQAEVDGEMVSLPHVEDAVVVLVCGAIAPLVLTQGAVPAP 241

QY 242 RHVVAAPRRDQIAGSSRTSVFFLRPNADFTFVPLARECGDVSLDGETATFQDWIGN 301
DB 242 NHVVSFDDASMLKSDRTSSVFFLRSTDTFTFSPDARKYGLDVSLDMKATFGDMIGTN 301

QY 302 YVNI 305
DB 302 YVTM 305

RESULT 5
US-09-413-231-5
; Sequence 5, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-413-231-9
Query Match 10.0%; Score 164; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QQGLHODEFRCLRDKGLFYLTDCGLTDTLTELKSAKDLVIDFEH 58
DB 10 VPTIDISPLFGTDAAAKRAVAEIIHGACRGSGFFYATNHGV---DVQQLQDVVNEFHGM 66

QY 59 SEAEK-----RAVTSVPVPMRRGFT-----TGLESESTAOITNTGSDYSCMYSGMT---A 106
DB 67 TDQEKHDLAIHAYNPDPNPHVRNGYYKAVPGRKAVESFCVYNLPDFGDHPM-IAAGTPMHE 125

QY 107 DNLFPSSG-----FGRIWTQYFDRQYTTASRAVAREVLRTATGTEPDGGVEA-----FLDCE 156
DB 126 VNLWPDDEERHPRFRPCGEGYRQMLKSLTVLMRGLALALG-RPEHFFDAALAEQDSLSV 184

QY 157 PLLRFPYQVPEHRSAAE-QPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAFD 215
DB 185 SLIRYPLEEYPPVKTPGPDQLLSFRDLDSVMTITLVFQTQVN-----LQVETVDGWRDI 240

QY 216 PYRPDAVLVFCGAIATLVTGGQVKAPRHVVAAPRRDQIAGSSRTSVFFLRPNADFTFV 275
DB 241 PTSENDFLVNCGTYMAHVTDNYFPAPNHRV-----KFVNAERLSLPFFLNGGHEAVIE- 293

QY 276 PLARECGDVSLDGETATFQDWI 298
DB 294 PFVPE-GASEEVNREALSYGDYL 315

RESULT 6
US-09-413-231-9
; Sequence 9, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus strain NRRL 3585
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-413-231-9
Query Match 10.0%; Score 164; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QQGLHODEFRCLRDKGLFYLTDCGLTDTLTELKSAKDLVIDFEH 58
DB 10 VPTIDISPLFGTDAAAKRAVAEIIHGACRGSGFFYATNHGV---DVQQLQDVVNEFHGM 66

QY 59 SEAEK-----RAVTSVPVPMRRGFT-----TGLESESTAOITNTGSDYSCMYSGMT---A 106
DB 67 TDQEKHDLAIHAYNPDPNPHVRNGYYKAVPGRKAVESFCVYNLPDFGDHPM-IAAGTPMHE 125

QY 107 DNLFPSSG-----FGRIWTQYFDRQYTTASRAVAREVLRTATGTEPDGGVEA-----FLDCE 156
DB 126 VNLWPDDEERHPRFRPCGEGYRQMLKSLTVLMRGLALALG-RPEHFFDAALAEQDSLSV 184
```

Db 126 VNLWDPDEERHPRPRPCEGYRQMLKSLTVLMRGLALG-RPEHFFDAALAEQDSSV 184
QY 157 PLLRERYFPQVPEHRSABE-QPLRMAPHYDLSMTLIIQOTPCANGFVSLQAEVGGFTDL 215
Db 185 SLIRYPYLSEYPPVKTPGQQLLSFRDLDSMTLVLFQTVQN-----LQVETVDGWRDI 240
QY 216 PYRPDAVLVFCGAIALTVTGQVKAPRHHVAAPRRDQIAGSRTSSVFFLRPNADFTFSV 275
Db 241 PISENDPLVNCGYMAHVINDYPPAHRV-----KFVNAERLSLPFFLNGGHEAVIE- 293
QY 276 PLARECGFVSLDGETATFQDWI 298
Db 294 PFVPE-GASEVNEALSXYGDL 315

RESULT 7

US-08-379-556A-10
; Sequence 10, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-10

Query Match 9.9%; Score 162; DB 2; Length 335;

Best Local Similarity 22.8%; Pred. No. 7.3e-10;

Matches 72; Conservative 55; Mismatches 135; Indels 54; Gaps 14;

QY 5 VPTFSLAELOOGLHQHDEFRCLDKGLFYLTDGCLTDTLKSADLVIDFEEHGSAEKR 64
Db 43 VPIDUSQADNESLVALISKASKDWGIFQVNHGIPSELISKLVNQVKGFEFFLPQ-EKE 101
QY 65 AVTSPVPMRRGPTGLESESTAQITNTGYSYDSCYMGTDANLF-----PSGDFGRW 119
Db 102 VIAPK-----DGKGYE-----CYGTLQKEVQKKGWDHLEHIVWPSFINYQW 148
QY 120 TQ-----YFD--ROYTASR-AVAREVL-----RATGTEPD-----GVFAFLDCEPLLR 162
Db 149 PKNPPSYRDTNBEYTSQSLGVANKLLGLLSKGLGLEBEDEVKQALGGEDLIY-----MLKIN 204

QY 163 YFPQVPEHRSABEQLRMAPHYDLSMTLIIQOTPCANGFVSLQAEVGGFTDLPYRPDAV 222
Db 205 YTPPCP-----CPALGVAPHTDMSSITL-----VPEVQGLQVFKDGQWDVAYIPNAL 256
QY 223 LVFCCGAIALTVTGQVKAPRHHVAAPRRDQIAGSRTSSVFFLRPNADFTFSVPLARECG 282
Db 257 ITHIGDQIEILSNGYKSYVHRSTVKNK-----KTRMSWPAFLPEPPEFVG-PIPKLVN 310
QY 283 FVSLDGETATFQDWI 298
Db 311 KODPPKYTKKYKDYV 326

RESULT 8

US-08-379-556A-8
; Sequence 8, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-8

Query Match 9.1%; Score 149; DB 2; Length 333;

Best Local Similarity 21.6%; Pred. No. 2.2e-08;

Matches 71; Conservative 57; Mismatches 126; Indels 74; Gaps 15;

QY 1 MDTT--VPTFSLAELOOGLHQHDEFRCL-----RDKGLFYLTDGCLTDTLKSADLVIDF 54
Db 39 LDTLVLEPAIDLS-----LEEDDVVKLVLSASKENGLFOVTNHGIPTEVIEKLVKVGKMF 93
QY 55 FPHGSEAEKRAVTSVPVPMRRGPTGLESESTAQITNTGYSYDSCYMGTDANLPPSGD 114
Db 94 FRAPAE-EKETAKPEGGVEGYGTMLQKEIQGR-----KGVWDHL----- 132
QY 115 FGRINT-----QYF-----DROYTASRAVAREVL-----RATGTEPDG-----GV 149
Db 133 FHKVMPSPSVINTQWPKPTSYREANETKYLRIVADKLFKCMKSLGLEBEDEVKSCGN 192
QY 150 EAPLDCPEPLLRPRYFPQVPEHRSABEQLRMAPHYDLSMTLIIQOTPCANGFVSLQAEVG 209

Db 193 E---DIVYLLKINYPCCPR-----PDALGVAHTDLSVITIL-----VPNDVAGLQVSRD 241
QY 210 GAFTDLPYRPDAVLVFCGAIATLVGTGQVKAPRHVAAPRRDQIAGSRTSSVFFLRPNA 269
Db 242 GRWYDVKIYIPNALIIHVGDOMEIMSGEYKAVLHRSTVNKE-----RTRISWPVFLPEPS 296
QY 270 DFTFSVPLARECGFVSLGSETATFOOM 297
Db 297 DFAVG-PIPKLISDERKPAKYKTKVFSEY 323

RESULT 9
US-09-413-231-10
; Sequence 10, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; NAME/KEY: MUTAGEN
; LOCATION: (214)
; OTHER INFORMATION: Glu214 in native IPNS modified to Arg
US-09-413-231-10

Query Match 8.7%; Score 143; DB 3; Length 333;
Best Local Similarity 23.2%; Pred. No. 1.1e-07;
Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;
QY 5 VPTFSLAEIQQGLHQD-----EPRCLRDKGLFYLTDGCLTDTLTKSAKDLVIDFFEH 57
Db 10 VPTIDISPL-FGTDPAKAHVAQINEACGSGFFVASHHG1---DVRRLQDVVNEFHT 65
QY 58 GSAEK-----RAVTSVPVPMRRGFTGLESESTAQITNTGSDYSNVCY---SMGTADNL 109
Db 66 MTQEKHDLAIHAYNENNSHVRNGY-----YMARPGKTVESWCYLNPSFGSDHPM 116
QY 110 FPSG-----DFGRITWQYFDROYTASRA-VAREVLRTGCT-----EPDG 147
Db 117 IKAGTWMHENVNMPDERHPDRSFGQYTYREVRFLSKVLLLLGFALALGKPEFFENEV 176
QY 148 GVEAFDLCPL-LRFYFPQVPE---HRSABEOLPMAPHYLSMTVLIQOTPCANGFVS 203
Db 177 TEEDTLSCRLMIRYPLDYPENAIKTPDGTFLSPRDLHDSMITVLFQTEVQN---- 232
QY 204 LQAEVGAFDLPYRPDAVLVFCGAIATLVGTGQVKAPRHVAAPRRDQIAGSRTSSVF 263
Db 233 LQVETVDGWSQSLPTSGENFLNGCTVGLYLTNDYFPAPNHRV-----KYVNAERLSLPF 286
QY 264 FL 265
Db 287 FL 288

RESULT 10
US-08-379-556A-2
; Sequence 2, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:

; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-379-556A-2

Query Match 8.6%; Score 140; DB 2; Length 347;
Best Local Similarity 22.3%; Pred. No. 2.5e-07;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;
QY 27 RKGLPYLTDGLTDTLTKSAKDLVIDFFEHSEAEKRAVT-SPVPTMERGF-TGLESES 84
Db 77 KEMGFQLINHIGIPDEADIADLQVKGKPEFHVHQEELIAKTPGSDIEGYTSLQKEV 136
QY 85 TQAITNTGSDYSVMCSMGTDNLF-----PSGDFGR1W-----TOYFDRQYT 128
Db 137 ESK-----KGWVDHLPHK1WPPSAVNYRYWPKNPPSYRANEYIGRMRE 181
QY 129 ASRAVAREVLRATGTEPDGGEVAFDCE--PLLRFYFPQVPEHRSABEQPLRMAPHYDL 186
Db 182 VVDRIFKSLSLGLGLEGHMEIEAAGGDEIVYLLKINYPCCPR----PDALGVVAHTDM 237
QY 187 SMTLIQOTPCANGFVSLQAEVGGAFDLPYRPDAVLVFCGAIATLVGTGQVKAPRHVA 246
Db 238 SVITIL-----VPEVQGLQVFDKGHWYDVKYIPNALIVHIGDQVEILSGKYSVYHRT 293
QY 247 APRDQIAGSRTSSVFFLRPNAD 270
Db 294 VNK-----DKTRMSWPFVLEPPSE 312

RESULT 11
US-09-413-231-8
; Sequence 8, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate

FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 321
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
OTHER INFORMATION: from Streptomyces cattleya
NAME/KEY: MUTAGEN
LOCATION: (211)
OTHER INFORMATION: Glu211 in native IPNS modified to Arg
US-09-413-231-8

Query Match 8.4%; Score 137; DB 3; Length 321;
Best Local Similarity 23.1%; Pred. No. 4.8e-07;
Matches 73; Conservative 32; Mismatches 149; Indels 62; Gaps 12;

QY 5 VPTFSLAELOOGLHQDEFR-----RCURDKGLFYLTDGLTDTLTKSAKDLVIDFFEH 57
DB 10 VPTIDISPOLFGDPTFRKTSRGRSTRPARGSGFFYASHHGIDVRRLOTWSN-----61
QY 58 GSEAEKRAVTSPPVTMRRTGFTGLSESTAQ-----ITNTGSDYSYMCY---SMGTADNLF 110
DB 62 -----ESTTMDQDSITWRSTRYNNENSHVNGYMARPGRETVESWCYLNPSFGEDHPWM 117
QY 111 PSG-----DFGRIWTOYDFRQVYASRAV---AREVLRTATGTEPDGGV-- 149
DB 118 KAGTPMHEVNVWPDEERHPDFGSGFGQY-HREVSASRRCGASRWRQAGESSNEVTE 176
QY 150 EARLDCPELLRFYFPQVPE---HRSABEQPLRMAPHYLSMVTLLQOTPCANGFVSLQA 206
DB 177 EDTLSAVSMIRYPYLPDYPPEAAIKTGPDRLSFRDHLDSMTILSKTEVQN---LQV 232
QY 207 EVGGATFDLPYRDAVLVFCGATATLVGGQVKAPRRHHVAAPRRDOIAGSRTSSVFFELR 266
DB 233 ETVDGHSQSLTSGENFLINGCTYGLTNDYFPAPNHRV-----KVNVAERLSLPFFLH 286
QY 267 PNADTFVSPLARECG 282
DB 287 AGQNSVMK-PFTRRTG 301

RESULT 12
US-09-413-231-4
Sequence 4, Application US/09413231
Patent No. 6284483
GENERAL INFORMATION:
APPLICANT: Dilley, David R
APPLICANT: Kadyrzhanova, Dina K
APPLICANT: Wang, Zhenyong
APPLICANT: Warner, Toni M
TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and Cephalosporins Under the Control of Bicarbonate
FILE REFERENCE: MSU41-453
CURRENT APPLICATION NUMBER: US/09/413,231
CURRENT FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 331
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
OTHER INFORMATION: from Penicillium chrysogenum
NAME/KEY: MUTAGEN
LOCATION: (212)
OTHER INFORMATION: Glu212 in native IPNS modified to Arg
US-09-413-231-4

Query Match 8.2%; Score 133.5; DB 3; Length 331;
Best Local Similarity 24.1%; Pred. No. 1.3e-06;
Matches 77; Conservative 39; Mismatches 131; Indels 73; Gaps 14;

QY 27 RDKGLFYLTDG-----LTDTELKSAKDLVIDFF--EHGSE-----60
DB 37 RDTGFFYAVNVHGVGVKRLSNKTRPHEFSITDE---KWDLAIRAYNKEHQDQIRAGYYLS 93
QY 61 -AEKAVTSPVPTMERGFTGLSESTAQITNTGTSYSDYSCYCMGTAD--NLFPS-----G 113
DB 94 IPEKKAVES-----FCYL-----NPNFKPDHPLIOSKTPTHEVNVNWPDEKHP 136
QY 114 DFGRIWTOYDFRQVYASRAVAREVLRTATGTEPDGGVEAF-----LDCEPLLRFRYFPQVP 168
DB 137 GFREFAEQYVWDFGLSSALLAGYALGKEBDFSRHFKKEDALSSVVLIRYPYLNPIP 196
QY 169 EH--RSABE-QPLRMAPHYDLSMVTLLQOTPCANGFVSLQAQVGGATFDLPYRDAVLV 225
DB 197 PAAIKTAEDGTKLSPRWHDVSLITVLYQSDVAN---LQVEMPOGYLDIEADNAYLVN 252
QY 226 CGAIATLVTTGGQVKAPRRHHVAAPRRDOIAGSRTSSVFLRDNADFTESVPLARECGFDV 285
DB 253 CGSYMAHITNNYYPAPIHRV-----KWNNEERQSLPFFV---NLGNFDTVQWDPSEKDG 304
QY 286 SLDGETATPDQWIGGNVYNI 305
DB 305 KTDQRPISYGDYLNQGLVSL 324

RESULT 13
US-08-553-367A-6
Sequence 6, Application US/08553367A
Patent No. 5939539
GENERAL INFORMATION:
APPLICANT: Theodor LANGE et al.
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,367A
FILING DATE: No. 5939539ember 27, 1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 49/FD.4.5M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO


```
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: zu PAT2353
; US-08-553-367A-6

Query Match      8.0%; Score 131, DB 2; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLPYLTDCGLTDELKSAKDLVIDFEHSGAEKRAVTSVPPT--MRRGFTGLESES--- 84
DB 89 GFPLVNHGVSESLIADAHRLMESFDMPLAGKQAKRPGESGCVASSFTGRFSTKLPW 148
QY 85 ----TAQITNTGYSYDSCYSGMTADNLPSPGDFGRITQYFDROYTASRAVAREVLRA 140
DB 149 KETLSQFQNDNSGSRVTQDYSDTLGQEF--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTETPDGGVEAFDCEPLLRFRYP--QVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
DB 207 LGVNRDYFRGFPEENDSINRLNHYPCQTP-----DLTLGTGPHCDPSSLTILHQDH-V 259
QY 199 NGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKAPRHVAAAPRRDQIAGSSR 258
DB 260 NG---LQVFVDNQWQSIIRPNKAFVNVIGDTFMALSGNGIFKSCILHRAVVNRE-----SAR 311
QY 259 TSSVFLRPNADFTFSVP 276
DB 312 KSWAFFLCPKQKVKVPP 329

RESULT 14
US-09-295-306-6
; Sequence 6, Application US/09295306
; Patent No. 6198021
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/295,306
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/553,367
; FILING DATE: No. 6198021ember 27, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 49/DIV-PD4.5M2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: zu PAT2353
; US-09-295-306-6

Query Match      8.0%; Score 131, DB 3; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLPYLTDCGLTDELKSAKDLVIDFEHSGAEKRAVTSVPPT--MRRGFTGLESES--- 84
DB 89 GFPLVNHGVSESLIADAHRLMESFDMPLAGKQAKRPGESGCVASSFTGRFSTKLPW 148
QY 85 ----TAQITNTGYSYDSCYSGMTADNLPSPGDFGRITQYFDROYTASRAVAREVLRA 140
DB 149 KETLSQFQNDNSGSRVTQDYSDTLGQEF--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTETPDGGVEAFDCEPLLRFRYP--QVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
DB 207 LGVNRDYFRGFPEENDSINRLNHYPCQTP-----DLTLGTGPHCDPSSLTILHQDH-V 259
QY 199 NGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKAPRHVAAAPRRDQIAGSSR 258
DB 260 NG---LQVFVDNQWQSIIRPNKAFVNVIGDTFMALSGNGIFKSCILHRAVVNRE-----SAR 311
QY 259 TSSVFLRPNADFTFSVP 276
DB 312 KSWAFFLCPKQKVKVPP 329

RESULT 15
US-09-734-719-6
; Sequence 6, Application US/09734719
; Patent No. 6455675
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/734,719
; FILING DATE: 13-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/295,306
; FILING DATE: April 21, 1999
; APPLICATION NUMBER: 08/553,367
; FILING DATE: No. 6455675ember 27, 1995
; APPLICATION NUMBER: PCT/EP94/01664
; FILING DATE: May 24, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 2000-1678/LC/01784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX: <Unknown>
```

```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 378 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;     ORGANISM: Arabidopsis thaliana
;     IMMEDIATE SOURCE:
;     CLONE: zu PAT2353
;   SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-734-719-6

Query Match      8.0%; Score 131; DB 4; Length 378;
Best Local Similarity 23.3%; Pred. No. 3e-06;
Matches 60; Conservative 40; Mismatches 130; Indels 28; Gaps 8;

QY 30 GLFYLTDGGLTDELKSAKDLVIDFPEHGSEAEKRAVTSVPVT--MRGFTGLESES--- 84
DB 89 GFPLVNVHGVSESLIADAHRLMESFFDMLAGKQAKRKPGEKSGYASSFTGRFSTKLEW 148
QY 85 ----TAQITNTGSGSYSCMYSMGTADNLFPSSGDFGR1WTQYFDROYTASRAVAREVLRA 140
DB 149 KETLSFQFSNDNSGSRVQDYFSDTLGQEF--QFGKVYQDYCEAMSSLSLKIMELLGLS 206
QY 141 TGTEPDGGVEAFIDCEPLARFRYP--QVPEHRSABEEQLRMAPHYDLSMVTLIQOTPCA 198
DB 207 LGVNRDYFRGFPEENDSIMRLNHYPPCQTP-----DLTLGTGFCHDCPSSLTILHODH-V 259
QY 199 NGFVSLQAEVGGFTDLPYRPPDAVLVFCGAIATLVGGQVKAPRHHVAAPRRDQIAGSSR 258
DB 260 NG---LQVFVDNQWQSIKRNPKAFVNVNIGDTFMALNSNGIFKSCSLHRAVVNRE-----SAR 311
QY 259 TSSVFFLRPNADFTFSVP 276
DB 312 KSMAFFLCPKKDKVVKPP 329
```

Search completed: February 3, 2004, 17:43:00
Job time : 22 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1624	99.2	311	10	US-09-924-841-18	Sequence 18, Appl
2	168	10.3	366	12	US-10-369-493-12239	Sequence 12239, A
3	164	10.0	329	10	US-09-924-841-5	Sequence 5, Appli
4	164	10.0	329	10	US-09-924-841-5	Sequence 5, Appli
5	148.5	9.1	327	12	US-10-369-493-14043	Sequence 14043, A
6	143	8.7	333	10	US-09-924-841-10	Sequence 10, Appl
7	140	8.6	359	15	US-10-136-444-8	Sequence 8, Appli
8	139.5	8.5	380	12	US-10-431-373-88	Sequence 88, Appl
9	137	8.4	321	10	US-09-924-841-8	Sequence 8, Appli
10	136.5	8.3	259	12	US-10-369-493-17939	Sequence 17939, A
11	133.5	8.2	331	10	US-09-924-841-4	Sequence 4, Appli
12	128.5	7.8	309	12	US-10-369-493-7649	Sequence 7649, Ap
13	128	7.8	320	12	US-10-369-493-14020	Sequence 14020, A
14	125	7.6	331	12	US-10-369-493-12421	Sequence 12421, A
15	124.5	7.6	317	12	US-10-369-493-4891	Sequence 4891, Ap

Qy	61	AEKRAVTSVPVPMRGGFTGLSESTAQITNTGSDYSYCMYSMCTADNLPSPSGDGRINT	120
Db	61	AEKRAVTSVPVPMRGGFTGLSESTAQITNTGSDYSYCMYSMCTADNLPSPSGDGRINT	120
Qy	121	QYFDRQYTSRAVAREVLRTGTETPDGVEAFLLDCEPLLRFYFPQVPEHRSAAEQPLRM	180
Db	121	QYFDRQYTSRAVAREVLRTGTETPDGVEAFLLDCEPLLRFYFPQVPEHRSAAEQPLRM	180
Qy	181	APHYDLSMWTLIQOTPCANGFVSQAEVGGAFDTLPYRPDAVLVFCGAIATLVTGGQVKA	240
Db	181	RPHYDLSMWTLIQOTPCANGFVSQAEVGGAFDTLPYRPDAVLVFCGAIATLVTGGQVKA	240
Qy	241	PRHHVAAPRDDQIAGSSRTSSVFFFLRPNADFTFSVPLARECGDPVSLDGETATFQDWIGG	300
Db	241	PRHHVAAPRDDQIAGSSRTSSVFFFLRPNADFTFSVPLARECGDPVSLDGETATFQDWIGG	300
Qy	301	NYVNIIRRTSKA	311
Db	301	NYVNIIRRTSKA	311

RESULT 2

US-10-369-493-12239
 ; Sequence 12239, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 12239
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Mesorhizobium loti
 US-10-369-493-12239

Query Match	Score	DB	Length
Best Local Similarity	10.3%	168;	366;
Matches	24.7%	Pred. No. 9.6e-10;	
85; Conservative	41;	Mismatches 158;	
Indels	60;	Gaps	14;

Qy	1	MOTTVTFSLAELOOGLHQ-----DEPRRCURDKGLFYLTDCGLUTDELKSAKDVIDRF	55
Db	28	MPRIVPLVLSRLAQAGSERRTFLADLSASRDIGFFYLAGHGHSWAEISEVLITASRQGF	87
Qy	56	EHGSEAEKAVTSPVPTWMRGFTGLESTTAQITNTGYSYSCYSGMTADNLFPSGDF	115
Db	88	AL-PEADKLAIEMWKSSQFRGYRAGBLTK-----GREDRERQLDIGVERQAIAGPG	140
Qy	116	GRIWTO-YFDRQVITAS-----RAVAREVLRTATGTEPPDGGVEAFDCEPLL	159
Db	141	TPAWTLOGPNQWPAALPDLKALLAWQSKYTAVAIRLLKAFQSLDQPSDAF---DPI-	196
Qy	160	RFRYPFQVPEHR-SAEEOPLR-----MAPHYDLSMVTLLIQOTPCANGFVSLQAEVGG	210
Db	197	----YSSEPNNRMKIVRYPGRTDTGGDQGVGAHKDGGFLTLQLQ-----DDNKGQLQVDYDG	248
Qy	211	AFDTLPRYDAVLVECGATATVTCGVQVKAAPHVHAAPRRDQTAGSSRTSVSFLRPND	270
Db	249	SWVDVDPIGTLLVNIIGELLEASNGYLURATVHRVTVP-----AGVERISVPVFFSARLD	304
Qy	271	FT-----FSVPLARECGFVDSLDGETATFQDWIGNGVYNNRTS	309
Db	305	ATIPPLGLSEELAAQAARGPAS-DPNQPLFRD-VGNTVLKSLRIS	346

RESULT 3

```

US-09-924-841-5
; Sequence 5, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-5

```

Query Match	10.0%	Score	164;	DB	10;	Length	329;
Best Local Similarity	23.8%	Pred.	No. 2.3e-09;				
Matches	77;	Conservative	50;	Mismatches	150;	Indels	46;
						Gaps	14;

Qy	5	VPTFSLAEI-----OQGLHQDEFRRLDKGFYLTDGCLTDTLTKLSAKDLVIDRFEHG	58
Db	10	VPTIDISPLFGTDAAKRVAEIHGACRGSGFFYATNHGV---DVQQLQDVVNEFHGAM	66
Qy	59	SEAEK-----RAVTSVPVMTRRGF-----TGLESESTAQITNTGYSYDSCYCSMGT---A	106
Db	67	TDQEKHDLAIHAYNPDPNPHVRNGYKAVPGKAVESFCYLNPDGFGDHPM-IAAGTPMHE	125
Qy	107	DNLFPSGD----PGRITWTQYDFDRQYTASRAVAREVLRATGETPDGGVEA-----FLDCE	156
Db	126	VNLWPDDEERHPRFRPECEGYRQMLKSTVLMRGLALALG-RPEHFHFDAAALAEQDLSVV	184
Qy	157	PLLRFRYFQVQVHRSAEE--QPLRMAPHYDLMSVTLIQOTPCANGFVSLQAEVCGAGTDL	215
Db	185	SLRYPYLEEYPVKTPGDGQLLSFRDHLDSWITVLFTQVQN---LQVETVDGWRDI	240
Qy	216	PYRPDAVLVPCGAIATLVTTGGQGVKAPRHHVAAPRRDQIAGSSRSTSSVFFLRPNADFTSV	275
Db	241	PTSENDFLVNCGYMAHVINDYFPANHRV-----KFVNAERLSLFFFLNGGHEAVIE-	293
Qy	276	PLARECGFDVSLDGETATFQDWI	298
Db	294	PFVPE--GAGEEVRNEALSYGDYL	315

RESULT 4

```

US-0924-841-9
; Sequence 9, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadvyrzhanova, Dina K
; APPLICANT: Wang Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841

```

```

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus strain NRRL 3585
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-9

Query Match      10.0%; Score 164; DB 10; Length 329;
Best Local Similarity 23.8%; Pred. No. 2.3e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QQGLHQDFRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFFEHG 58
Db 10 VPTIDISPLPGTDAAKRVAEIHGACRSGGFYATNRGV---DVQQLQDVVNEFHGAM 66

QY 59 SEAEK-----RAYTSPVPTWRRGF---TGLESESTAQITNTGSYSYDSCYSGMT---A 106
Db 67 TDQEKHDLAIHAYNPDPNPHVNGYKAVPCRKAVESFCYLNLPDGEDHPM-IAAGTPMHE 125

QY 107 DNUFPSPGD-----FGRITWQYDFRQYATASRAVAREVLRAATGTEPDGGVEA-----FLDCE 156
Db 126 VNLWPDDEERHPRFPCEGYRQMLKSLSTVLMRGLALALG-RPEHFFDALAEQDSLSSV 184

QY 157 PLLRERYFPQVPEHRSABE-QPLRMAPHYDLSMTLIQQTCCANGFVSLQAEVGGAFD 215
Db 185 SLRYPYLEEYPPVKTPGPDQLLSFRDLQVSMITVLFQVQN-----LQVETVQGRDI 240

QY 216 PYRPDAVLVFCGAIATLVTTGGQVKAPRRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
Db 241 PTSENDVLNCGTYMAHVNTDYPAPNHRV-----KFVNAERLSLPFLGNGGHEAVIE- 293

QY 276 PLARECGFVSLDGETATFODWI 298
Db 294 PFVPE-GASEEVRNEALS YGDYL 315

RESULT 5
US-10-369-493-14043
; Sequence 14043, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14043
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-14043

Query Match      9.1%; Score 148.5; DB 12; Length 327;
Best Local Similarity 23.8%; Pred. No. 1.2e-07;
Matches 80; Conservative 43; Mismatches 162; Indels 51; Gaps 13;

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus strain NRRL 3585
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-9

Query Match      10.0%; Score 164; DB 10; Length 329;
Best Local Similarity 23.8%; Pred. No. 2.3e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QQGLHQDFRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFFEHG 58
Db 10 VPTIDISPLPGTDAAKRVAEIHGACRSGGFYATNRGV---DVQQLQDVVNEFHGAM 66

QY 59 SEAEK-----RAYTSPVPTWRRGF---TGLESESTAQITNTGSYSYDSCYSGMT---A 106
Db 67 TDQEKHDLAIHAYNPDPNPHVNGYKAVPCRKAVESFCYLNLPDGEDHPM-IAAGTPMHE 125

QY 107 DNUFPSPGD-----FGRITWQYDFRQYATASRAVAREVLRAATGTEPDGGVEA-----FLDCE 156
Db 126 VNLWPDDEERHPRFPCEGYRQMLKSLSTVLMRGLALALG-RPEHFFDALAEQDSLSSV 184

QY 157 PLLRERYFPQVPEHRSABE-QPLRMAPHYDLSMTLIQQTCCANGFVSLQAEVGGAFD 215
Db 185 SLRYPYLEEYPPVKTPGPDQLLSFRDLQVSMITVLFQVQN-----LQVETVQGRDI 240

QY 216 PYRPDAVLVFCGAIATLVTTGGQVKAPRRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV 275
Db 241 PTSENDVLNCGTYMAHVNTDYPAPNHRV-----KFVNAERLSLPFLGNGGHEAVIE- 293

QY 276 PLARECGFVSLDGETATFODWI 298
Db 294 PFVPE-GASEEVRNEALS YGDYL 315

RESULT 5
US-10-369-493-14043
; Sequence 14043, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14043
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-14043

Query Match      9.1%; Score 148.5; DB 12; Length 327;
Best Local Similarity 23.8%; Pred. No. 1.2e-07;
Matches 80; Conservative 43; Mismatches 162; Indels 51; Gaps 13;

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus strain NRRL 3585
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-9

Query Match      10.0%; Score 164; DB 10; Length 329;
Best Local Similarity 23.8%; Pred. No. 2.3e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAEL-----QQGLHQDFRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFFEHG 59
Db 1 LPILDLSLLDGTSPSQROAFLDRLHAARDVGFYLTGHTGIDAGLLREQVDYARQFAL-P 59

QY 60 EAEKRAVTSFVPTMRGFTGLESESTAQITNTGSYSYDSCYSGMTADNLFPSPGDFGRW 119
Db 60 DSEKNAVGMINSFHRGYNRAASEITR-----GQPDQREQFDLGAERDVLPLNADSP 113

QY 120 TO-----YFDRQYATASRAVAREVLRAATG---TEPDGGVEAFDCEPLL 159
Db 114 ARLOGNPQWPGALPQLKPLLLDQWQWNTOMSLR-LLRAFAQALSLEDAFDRLYGDKPNE 172

QY 160 RFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQQTCCANGFVSLQAEV-GGAFTDLPYR 218
Db 173 HIKLMRYPGQASTASNQ--GUGAHKDSGFLSFLQDQQA---GLQVEIEEGRWIDALPR 226

QY 219 PRAVLVFCGAIATLVTTGGQVKAPRRHHVAAPRRDQIAGSSRTSSVFFLRPNAD-----FTF 273
Db 227 DNTLVVNIIGELLELATNGYLRAIVHRVSP-----VGSERLSIAFFLGAQLDVAVPLPL 282

QY 274 SVPLARECGFVSLDGETATFODWIGNVNIRRTS 309
Db 283 PTALLREARCPAS-DPLNPLFRD-VGVNLYKGLRS 316

RESULT 6
US-09-924-841-10
; Sequence 10, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Tom M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from streptomyces anulatus ( S. lipmanii)
; NAME/KEY: MUTAGEN
; LOCATION: (214)
; OTHER INFORMATION: Glu214 in native IPNS modified to Arg
US-09-924-841-10

Query Match      8.7%; Score 143; DB 10; Length 333;
Best Local Similarity 23.2%; Pred. No. 5.1e-07;
Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;

QY 5 VPTFSLAELQQGLHQD-----EPRRCRLDKGLFYLTDCGLTDTTELKSAKDLVIDFFE 57
Db 10 VPTIDISPL-FGTDPAKAKHVARQINEACRSGGFVYASHHGI---DVRRLQDVVNEPHRT 65

QY 58 GSEAEK-----RAYTSPVPTMRGFTGLESESTAQITNTGSYSYDSCY---SMGTADNL 109
Db 66 MTQDQEKHDLAIHAYNNNSHVRNGY-----YMARPGKTKVESWCYLNPSFGEDHPM 116

QY 110 FPSG-----DFGRITWQYDFRQYATASRA-VAREVLRAATG-----EPDG 147
Db 117 IKAGTPMHEVNWVWPDDEERHPRFSGEQYVREVRLSKULLRGPALALGKPEFFENEV 176

QY 148 GVEAFDCEPL-LRFRYFPQVPE---HRSAEQPLRMAPHYDLSMTLIQQTCCANGFVS 203
```

Db 177 TEBDTLSCRLMRYPLDYPPEAAIKTGPDGRLSFRDHLDSVMTVLFFQTEVQN----232
QY 204 LQEVGGAFTDLPRDVLVFCGATATLVGGQVKAAPRHVAAAPRRDQIAGSSRTSSVP 263
Db 233 LQVETVDGWSLPTSGENFLNGTGLVLTNDYFPAPNHRV-----KYVNAERLSLPP 286
QY 264 FL 265
Db 287 FL 288

RESULT 7
US-10-136-444-8
; Sequence 8, Application US/10136444
; Publication No. US20030101477A1
; GENERAL INFORMATION:
; APPLICANT: COLLIVER, Steven P.
; APPLICANT: HUGHES, Stephen G.
; APPLICANT: MUIR, Shelagh R.
; APPLICANT: TUNEN van, Adrianus J.
; APPLICANT: VERHOEYEN, Martine E.
; TITLE OF INVENTION: A PROCESS FOR INCREASING THE FLAVONOID CONTENT OF A PLANT AND PLA
; FILE REFERENCE: 056159-5106
; CURRENT APPLICATION NUMBER: US/10/136,444
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: EP 01304009.2
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Petunia hybrida
US-10-136-444-8

Query Match 8.6%; Score 140; DB 15; Length 359;
Best Local Similarity 22.3%; Pred. No. 1.2e-06;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;
QY 27 RDXGLFYLDCGLTDLTSLAKDLVIDDFEHGSAEAKRAVT-SPVPTMRGF-TGLESES 84
Db 88 KEWGIFQLNHGIPDEAIDLQKVGKEFFEHVQEEKELIAKTPGSDNDIEGYGTSIQKEV 147
QY 85 TAQITWGSYDSYCMYSGTADNLF-----PSGDFGRW-----TQYFDRQYT 128
Db 148 EKG-----KGVVDHLFKIWPSSAVNRYWPKNPPSYREANEYGRMRE 192
QY 129 ASRAVAREVLRAATGTPDGGVEAFDCE--PLLRFRYFPQVPEHRSABEQPLRMAPHYDL 186
Db 193 VVDRIKFSLSLGLGHEHIEAAGDEIVYLKINYPCCPR-----PDLALGVVAHTDM 248
QY 187 SMVTLLIQTPCANGFVSLQAEVGAFTDLPRDVLVFCGATATLVGGQVKAAPRHVAA 246
Db 249 SVITIL---VPNEVQLQVFKDGHVYDKYIPNALIVHIGDQVEILSNGKYKSVYHRTT 304
QY 247 APRDQIAGSSRTSSVFFLRPNAD 270
Db 305 VNK-----DKTRMSWPVLEPPSE 323

RESULT 8
US-10-431-273-88
; Sequence 88, Application US/10431273
; Publication No. US20030237108A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: Glenn, Matthew
; APPLICANT: No. US20030237108A1riss, Michael Geoffrey
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire

; APPLICANT: Forster, Richard L. S.
; TITLE OF INVENTION: Compositions isolated from forage
; FILE REFERENCE: 11000.10690
; CURRENT APPLICATION NUMBER: US/10/431,273
; CURRENT FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Lolium perenne
US-10-431-273-88

Query Match 8.5%; Score 139.5; DB 12; Length 380;
Best Local Similarity 23.3%; Pred. No. 1.5e-06;
Matches 70; Conservative 43; Mismatches 140; Indels 47; Gaps 11;
QY 1 MDTVTFTSLAELOQLHQDEPR-----RCLRDKGLFYLDCGLTDLTSLAKDLV 51
Db 60 LDTSPVIDLDELVAAT-ADEGRMRQIMEAVAAACREWGFFQVNVHGVAPELMHAREAW 118
QY 52 IDFFEHGSAEAKRAVTSVPVPTMRGFTGLESESTQITNTGYSYCMYSGTADNLP 111
Db 119 RGFRLPITAKQYANLP-----RTVEGYGSRVGVKGGPLDWGDY---YFLHLPADGK 170
QY 112 SGDFGRWITQYFDRQYTASRAVAREVL-----ATCTEPDGGVEAF--LDCEPL 158
Db 171 SPD--KYMPTNPAICDVSEYGREVIRCELLMKMSASLGLEATRFQEAFFGSGCVC 228
QY 159 LRFRYFPQVPEHRSABEQPLRMAPHYDLSMVTLLIQTPCANGFVSLQAEVGAFTDLPYR 218
Db 229 LRANYPPRCQ---PDLTLGLSAHSDPGVLTVLADAEHVRGLQVRRAD-GEWVTVQPAR 283
QY 219 PDVAVLFCGATATLVGGQVKAAPRHVAAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLA 278
Db 284 HDAFTVNVGDQIQLSLSNMYKSVEHVRVMVNAKEE-----RISLALFYNPRGD----VPIA 334

RESULT 9
US-09-924-841-8
; Sequence 8, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE REFERENCE: MS041-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces cattleya
; NAME/KEY: MUTAGEN
; LOCATION: (211)
; OTHER INFORMATION: Glu211 in native IPNS modified to Arg
US-09-924-841-8

Query Match 8.4%; Score 137; DB 10; Length 321;
Best Local Similarity 23.1%; Pred. No. 2.3e-06;
Matches 73; Conservative 32; Mismatches 149; Indels 62; Gaps 12;

Qy	5	VPTFSLAELOOQLHODEFR-----RCLRKGLPYLTDGCLTDTTELKSAKDVLVIDFFEH	57
Dd	10	VPTIDISPOLFGTDPTRTSRGSRTPARGSGFFYASHHGIDVRRLLQTWSN-----	61
Qy	58	GSEAEKRAVTSVPVWRGFGTCLESESTAQ---ITWTGSYSDSYMCY---SMGTADNL	110
Dd	62	----ESTTMDQRSTTWSTRYNENNSHVRNGYMARPGRETVESWCYLNPFSGEODHPMM	117
Qy	111	PSG-----DFGRWTOFYFDQYTASRAV---AREVLARATGETPDGV--	149
Dd	118	KAGTMPHEVNVWPDERHPDFOGPEQY-HREVASRRCGCCASRWRRQAGESSNEVTE	176
Qy	150	EAFLDCEPLLRFRYPQVPE---HRSAEEQLRMAPHYDLMSVTLIQOTPCANGFVSQA	206
Dd	177	EDTL SAVSMIRYPYLPDYPEAAIKTGPDGTRLSPRDHLDVSMITVLSKTEVN---	232
Qy	207	EVGGAETDLPYRPDAVLVFCGAIATLVTTGGQVKAPRHVVAAARRDDQIAGSSRTSSVF	266
Dd	233	ETVDGQSLPTSGENFLNCGYLYLTNDYFPAPNHRV-----KYVNAERLSLPFFLH	286
Qy	267	PNAFTFSVPLARECG	282
Dd	287	AGONSVMK-PFTRRTG	301

RESULT 10
 US-10-369-493-17939
 ; Sequence 17939, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 17939
 ; LENGTH: 259
 ; TYPE: PRT
 ; ORGANISM: SPHINGOMONAS
 US-10-369-493-17939

Query Match	8.3%;	Score 136.5;	DB 12;	Length 259;
Best Local Similarity	25.2%;	Pred. No. 1.8e-06;		
Matches	67;	Conservative 41;	Mismatches 111;	Indels 47; Gaps 11;
Qy	30	GLFVLTGGLTDTLKSADLVIDPFPHGSAEAKVATSPVPTWRRGFTGLESESTAQIT	89	
Db	11	GFAVVRDHGVQEIETIAERWLTKAFFDL-PEEEKRAHFIPGGGGARGYTPPKTE----	65	
Qy	90	NTGSYSYSMCYSWG-----TADNLFPSGDFGRITWQYFDRQYATASAVAREV	137	
Db	66	KGATHDLKKEFWHIGROLAAGHRFADVMAPNIWPTREP--FRETFTIELFAAFOAGDKL	123	
Qy	138	LRAT-----GTPEPDGGVEAFLDCEPLLRFYFPQVPEHRSAEQQLRMAPHVDLGSVTLIQ	193	
Db	124	LSAVARYLGLAPDWFDTAVKDGNVLRLLHYPPV-----AADAPEVRAGAHEDINLIITLL-	178	
Qy	194	QTPCANGFVSLQAEVGG-AFTD-----LPYRP--DAVLVFCGAIAITLVGGQVKAPRHV	245	
Db	179	-----LGAEAGLELLDRGKMLAVKPEPGAMVINVDMLQRLTNHVLFSSTHVV	228	
Qy	246	AAPRRDIAGSSRTSSVFFFLRPNADF	271	
Db	229	VNPPAER-RGHSRYSMPPFFLHPADPF	253	

```

RESULT 11
US-09-924-841-4
; Sequence 4, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; FILE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPN
; OTHER INFORMATION: from Penicillium chrysogenum
; NAME/KEY: MUTAGEN
; LOCATION: (212)
; OTHER INFORMATION: Glu212 in native IPNS modified to Arg
; US-09-924-841-4

```

Query Match	8.2%;	Score 133.5;	DB 10;	Length 331;
Best Local Similarity	24.1%;	Pred. No. 5.8e-06;		
Matches	77;	Conservative 39;	Mismatches 131;	Indels 73; Gaps 14;
Qy	27	RDKGLFVLTDCG-----LDTTELKSAKDLVIDPF--EHGSE-----	60	
Db	37	RDTEGFFVANHGVVDVKELSKNTRFHFSTIDEE---KWDLAIRAYNEHQDQIRAGYLS	93	
Qy	61	-AEKRAVTSVPVMTMRGFTGLESESTAQITNTGYSYDSCYCSMGTD--NLFPF----	113	
Db	94	IPEKKAVER-----FCYL-----NPNPKPDHPLIQSKTPTHEVNVWPDEKHP	136	
Qy	114	DFGRHWTQYDFDRQYTSARAVAREVLRATGTEPDGCGVEAF-----LDCEPLLRFYFPQVP	168	
Db	137	GFREFAEQYWDVFGLESSALLRGYALGKBEEDFFSRHFHKEDALSSVILIRYPYLPNIP	196	
Qy	169	EH--RSABE-QPLRMAPHYLDLSSVMTLLQOQTPCANGFVSLQAEVGGATFDLPYRDDAVLAF	225	
Db	197	PAAIKTAEDGTKLSFRWHEDVSLITVLYQSDVAN-----LQVEMDQGYLDIDEADNNAYLVN	252	
Qy	226	CGAIIATLVITGGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDV	285	
Db	253	CGSYMAHITNNYYPAPIHRV-----KWVNEERQSLPFFV--NLGFNDITQVPWDPSEKD	304	
Qy	286	SLDGETATFDQDWIGGNNVNI	305	
Db	305	KTDQRPISYGDYQLONGLVSL	324	

```

RESULT 12
US-10-369-493-7649
; Sequence 7649, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Jiongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MIC
; TITLE OF INVENTION: PLANTS WITH IMPR
; FILE REFERENCE: 38-10(52052)B

```


Search completed: February 3, 2004, 17:42:26
Job time : 34 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:35:49 ; Search time 175 Seconds
(without alignments)
1617.056 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTVTFTSLAELOQLHQD.....ATFDWIGNYNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
- 32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1637	100.0	311	19	US-09-582-486-1 Sequence 1, Appli

2	1629	99.5	311	22	US-09-791-537-32088	Sequence 32088, A
3	1627	99.4	311	22	US-09-791-537-11094	Sequence 11094, A
4	1624	99.2	311	24	US-09-924-841-18	Sequence 18, Appl
5	1186.5	72.5	314	22	US-09-791-537-121286	Sequence 121286,
6	965.5	59.0	310	6	US-08-233-605-6	Sequence 6, Appli
7	965.5	59.0	310	11	US-08-727-468-6	Sequence 6, Appli
8	965.5	59.0	310	22	US-09-791-537-10913	Sequence 10913, A
9	939.5	57.4	318	22	US-09-791-537-10914	Sequence 10914, A
10	937	57.2	319	22	US-09-791-537-4795	Sequence 4795, Ap
11	926.5	56.6	332	22	US-09-791-537-94436	Sequence 94436, A
12	852	52.0	313	22	US-09-791-537-4796	Sequence 4796, Ap
13	179	10.9	329	22	US-09-791-537-89857	Sequence 89857, A
14	176	10.8	329	22	US-09-791-537-9085	Sequence 9085, Ap
15	168	10.3	366	29	US-10-369-493-12239	Sequence 12239, A
16	168	10.3	366	32	US-60-360-039-12239	Sequence 12239, A
17	164	10.0	329	22	US-09-791-537-89855	Sequence 89855, A
18	164	10.0	329	24	US-09-924-841-5	Sequence 5, Appli
19	164	10.0	329	24	US-09-924-841-9	Sequence 9, Appli
20	162	9.9	326	22	US-09-791-537-9082	Sequence 9082, Ap
21	162	9.9	326	22	US-09-791-537-89847	Sequence 89847, A
22	155	9.5	342	28	US-10-219-999-52994	Sequence 52994, A
23	155	9.5	375	30	US-10-425-114-58493	Sequence 58493, A
24	154.5	9.4	366	30	US-10-437-963-146275	Sequence 146275,
25	154	9.4	352	28	US-10-219-999-51128	Sequence 51128, A
26	154	9.4	352	30	US-10-425-114-63175	Sequence 63175, A
27	153	9.3	339	21	US-09-708-427-77410	Sequence 77410, A
28	153	9.3	342	21	US-09-708-427-77409	Sequence 77409, A
29	153	9.3	342	28	US-10-219-999-41198	Sequence 41198, A
30	153	9.3	342	28	US-10-219-999-59894	Sequence 59894, A
31	153	9.3	342	32	US-60-324-109-17598	Sequence 17598, A
32	153	9.3	350	30	US-10-437-963-136698	Sequence 136698,
33	153	9.3	371	28	US-10-219-999-52863	Sequence 52863, A
34	153	9.3	371	30	US-10-425-114-62424	Sequence 62424, A
35	153	9.3	372	30	US-10-425-114-53454	Sequence 53454, A
36	153	9.3	372	30	US-10-425-114-72059	Sequence 72059, A
37	151	9.2	331	22	US-09-791-537-166840	Sequence 166840,
38	150.5	9.2	338	21	US-09-708-427-50491	Sequence 50491, A
39	150.5	9.2	341	21	US-09-708-427-50490	Sequence 50490, A
40	149.5	9.1	327	24	US-09-902-540-15162	Sequence 15162, A
41	149	9.1	440	21	US-09-733-089-18920	Sequence 18920, A
42	149	9.1	440	23	US-09-816-660-18920	Sequence 18920, A
43	148.5	9.1	327	29	US-10-369-493-14043	Sequence 14043, A
44	148.5	9.1	327	32	US-60-360-039-14043	Sequence 14043, A
45	148.5	9.1	389	30	US-10-437-963-148589	Sequence 148589,

ALIGNMENTS

RESULT 1

- US-09-582-486-1
- Sequence 1, Application US/09582486
- GENERAL INFORMATION:
- APPLICANT: SCHOFIELD, Christopher J.
- APPLICANT: BALDWIN, Jack E.
- APPLICANT: LLOYD, Matthew D.
- APPLICANT: HARLOS, Karl
- APPLICANT: ANDERSSON, Inger
- APPLICANT: TERWISSCHA VAN SCHELTINGA, Anke S.
- APPLICANT: VALEGARD, Karin
- APPLICANT: RAMASWAMY, S.
- TITLE OF INVENTION: MODIFIED DEACETOXYCEPHALOSPORIN C SYNTHASE (DAOS) AND X-RAY STRUCTURE
- FILE REFERENCE: 08004624
- CURRENT APPLICATION NUMBER: US/09/582,486
- CURRENT FILING DATE: 2000-06-23
- PRIOR APPLICATION NUMBER: 9727370.0
- PRIOR FILING DATE: 1997-12-24
- PRIOR APPLICATION NUMBER: 9813644.3
- PRIOR FILING DATE: 1998-06-24
- NUMBER OF SEQ ID NOS: 1
- SOFTWARE: Patentin Ver. 2.1
- SEQ ID NO 1

PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 18
LENGTH: 311
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: modified DAOCs
OTHER INFORMATION: from Streptomyces clavuligerus
NAME/KEY: MUTAGEN
LOCATION: (181)
OTHER INFORMATION: Ala181 in native DAOCs modified to Arg
US-09-924-841-18

Query Match 99.2%; Score 1624; DB 24; Length 311;
Best Local Similarity 99.4%; Pred. No. 2.7e-177;
Matches 309; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 MDTTPTFSLAELOQGLHQBDEFRRCLRDKGLFYLTDCGLTDTLKSADLVIDFFHSGSE 60
DB 1 MDTTPTFSLAELOQGLHQBDEFRRCLRDKGLFYLTDCGLTDTLKSADLVIDFFHSGSE 60

QY 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120
DB 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120

QY 121 QYEDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180
DB 121 QYEDROYTASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAJATLVTCGVKA 240
DB 181 RPHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAJATLVTCGVKA 240

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300

QY 301 NYNIRRTSKA 311
DB 301 NYNIRRTSKA 311

RESULT 5
US-09-791-537-121286
; Sequence 121286, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 121286
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Streptomyces lactamdurane
US-09-791-537-121286

Query Match 72.5%; Score 1186.5; DB 22; Length 314;
Best Local Similarity 69.8%; Pred. No. 7.1e-127;
Matches 217; Conservative 41; Mismatches 52; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOQGLHQBDEFRRCLRDKGLFYLTDCGL-TDTLKSADLVIDFFHSGSE 60
DB 3 DATVPTFDLAELEGLHQBDEFRRCLREKGVFLYKGTGLPAEADHASGREIAVDFFDHGTE 62

QY 61 AEKRAVTSPTVTRRGFTGLESTTAQITNTGYSYDSCYSGMTADNLPSPGDFGRWT 120

DB 63 AEKKAVMTPITPIRRGYAGLESESTAQITNTGKYDYSMSGTADNLPSPAEFKAWE 122

QY 121 QYFDRQYASRAVAREVLRTATGTEPDGGVEAFDCEPLLRFRYPQVPEHRSAAEQPLRM 180

DB 123 DYFARMYRASQDVARQVLTSGAEPEVGMADFDCPEPLLRFRYPQVPEHRSAAEQPLRM 182

QY 181 APHYDLSMVTLIQOTPCANGFVSLQAEVGAFTDLPYRDPDAVLVFCGAJATLVTCGVKA 240

DB 183 APHYDLSIVTLIHQTPCANGFVSLQAEVVDGSDYDIPAPQCAVLVFCGAVATLVADGAIKA 242

QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFDQWIGG 300

DB 243 PKHVAAPGADKRVGSSRTSSVFFLRPNADFTFSVPLARECGFDVSPAETATFDQWIGG 302

QY 301 NYNIRRTSKA 311

DB 303 NYNIRRTSKA 313

RESULT 6
US-08-233-605-6
; Sequence 6, Application US/08233605
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan P.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enquita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/233,605
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-233-605-6

Query Match 59.0%; Score 965.5; DB 6; Length 310;
Best Local Similarity 61.2%; Pred. No. 2e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAELOQGLHQBDEFRRCLRDKGLFYLTDCGLTDTLKSADLVIDFFHSGSEA 61
DB 3 DKTVPFVMAELRDSGRQDEFREWAR-RGVFYLTYGATERDHRVATDAMDFFAQGTAE 61

Qy	62	EKGAVTISVPVPTWRRGFTGLESESTAQITNTGTSYSDYSWCMYSGMTADNLPPSGDFGRITWQ	121
Db	62	EKGAVTTKVPPTWRRGYSALEASTAQVNTGTYYTDSYSGMIGNGLPFSKEFESVWTD	121
Qy	122	YFDQVYTASRAVAREVLVLRATGTEPDGGVEAFDCEPLLRFPRYPFQVPEHRSABEQPLRMA	181
Db	122	YFDSLRYRAAQETARLVLTAACTYDGEDLDTLLDCDPVLRRLRYFPEVPEHRAAEYEPRRMA	181
Qy	182	PHYDLSMWTLIQOTPCANGFVLSQAEVGGAFDTDLPYRPDAVLVFCGAIATLVTGGGVKAP	241
Db	182	PHYDLSIIITFIHQTEPCANGFVLSQAEVDGEMVSLPHVEDAVVLCGAIAPLVTQGGAVPAP	241
Qy	242	RHHVAAPRBDQIAGSSRTSSVFFFLRPNADFTFESVPLARECGFVDSLDGETATFDQDWIGN	301
Db	242	NHHVSPDASMLKGSDRISSVFFLRPSTDTFFSVDPDARKYGLDVLDMKATFDGMDWIGN	301
Qy	302	YVNI	305
Db	302	YVTM	305

RESULT 7

```

US-08-727-468-6
; Sequence 6, Application US/08727468
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan P.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llarena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,468
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/233,605
; FILING DATE: 22-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-727-468-6

```

RESULT 8

```

US-09-791-537-10913
; Sequence 10913, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILIES
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10913
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Streptomyces lactamdurans
US-09-791-537-10913

Query Match          59.0%; Score 965.5; DB 22; Length 310;
Best Local Similarity 61.2%; Pred. No. 2e-101;
Matches 186; Conservative 37; Mismatches 80; Indels 1; Gaps 1;

Qy      2   DTTVPFSLAELOOGLHQHDEFRECLDKGLFYLTDCGLTDTCLKSAKDVLIDRPFHGESEA 61
Db      3   DKTVPVFSMAELRGDSGRQDEFREWAR -RGVFYLTGCGATERDHRVATDTAMDFFAQGTAE 61

Qy      62  EKRAVTSPPVPTMRGFTGLESESTAQITNTGYSYSDYSCMSYSGMTADNLFPSGDFGRITWQ 121
Db      62  EKQAVTKVPTMRGYSALAEASTAQVNTGTITDYMSYSMGIIGNLFPSEKFEESVWTD 121

Qy      122 YFRQVYTSARAVAREVLIRATGTPDGGVEAFDLCEPILLRPYPQVPEHRSABEQPLRMA 181
Db      122 YFDSLRYAAQETARLVLTAAAGTYDGEDDPLLLDCDPVLRILRYPFVPEHRAAAEYEPARMA 181

Qy      182 PHYDLNSMVTLLIQTPCANGFVSQAQEVGGAFTDLPYRPDAVLVFCGAIATLVTGSGQVKAP 241
Db      182 PHYDLGIITFIHQTPCANGFVSQAQEVGGMVSLPHVEDAVVULCGAIAPLVTQGAVPAP 241

Qy      242 RHHVAAPRRDQIAGSSRSTSSVFFLRPNADPTFFSVPLARECGFDVSLDGETATTFQDWIGGN 301
Db      242 NHHVSPDASMLKSGDRSTSSVFFLRPSTDTFFSVDPARKYGLDVSLDWEKATFGDWIGTN 301

Qy      302 YVNI 305
Db      302 YVTM 305

```

```

RESULT 9
US-09-791-537-10914
; Sequence 10914, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10914
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-791-537-10914

```

Query Match	57.4%	Score 939.5;	DB 22;	Length 318;
Best Local Similarity	58.6%	Pred. NO. 2.1e-98;		
Matches 180:	Conservative	43;	Mismatches 81;	Indels 3;
				Gaps 1;

Qy	2	DTTVPTTSLAELOOGLHODEFRRCURDKGLFVLTDCGLTDTTELKSAKOLVLDFFPHGSEA	61
Db	3	DTPVPIFNLAALREGAQOQEKRECVTGMGVPLTGYGAGDKDHLRLATLTAMDFFPANGTEA	62
Qy	62	EKRAVTSVPVPTMRRGFTGLESESTAQIINTQSYSDYSMCTADNLFPSGDGFGRIWTQ	121
Db	63	EKAAVTTDVPVPTMRRGYSALAEASTAQVTRTGSYTDYSMSFGISGNVFPSPPEPERVWTE	122
Qy	122	YFDROYTASRAVAREVLBATG---TEPPGGVEAFLDCEPLLRFRFPQVPEHRSAAEQPL	178
Db	123	YFDKLYAAAQETALRLVLTASGGYDAEI VGSJDELLDADFVLRRLRYFVPEPEHRSAAEHR	182
Qy	179	RMAPHYDLMSVTLTIQOTPCANGFVSLQAEVGGAFDTLFPYRDPDAVLVFCGAIATLVTGGV	238
Db	183	RMAPHYDLSIITFIHQTPCANGFVSLQAEI GGELVSLPVVEDAVVVMCGAMAPLATQAL	242
Qy	239	KAPRHVAA PRDQIAGSSRTSSVFFFLRPNADFTFSVPLARECGDFVSLDGETATFDQWI	298
Db	243	PAPRHVRSPCAGMREGSDRTSSVFFFLRPTTDFSFSAKARSYGLAVLDMETATFGDWI	302
Qy	299	GGNYVNI	305
Db	303	GTNYVTM	309

```

RESULT 10
US-09-791-537-4795
; Sequence 4795, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4795
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Lysobacter lactamgenus
US-09-791-537-4795

```

Query Match 57.2%; Score 937; DB 22; Length 319;
Best Local Similarity 58.0%; Pred. No. 4.1e-98;
Matches 177; Conservative 45; Mismatches 83; Indels 0; Gaps 0;

Qy	2	DTVTFTSLAELOOGLHODERRCLROKGLFYLTDCLGTTELKSAKDVLIDFPEHCSEA	61
Db	3	DSGIQIFDLELHGVRUDSPKSLFKGVFYVREDDSIKTEHAKAMDAVMDLFENGSAE	62
Qy	62	EKRAVTSVPVPTMRRGFTGLESESTAQITNTGYSYDSMYSMGTDNLFPSPGDGFRIMTQ	121
Db	63	QKNALRNLTNPVNRGFSDELEASTARITKCGEYDYSMWYSIGLITDNLFPSPAFAEIMTQ	122
Qy	122	YFDROYTASRAVAREVLKATGTEPDGGVEAFDCEPELLRFRYPPOVPEHRSASAEQPLRMA	181
Db	123	YFDRFYARTQDTRASVLRASDAGLDESVDLEFDCDPLLRFRFPEVEPEDRCARQEPERMA	182
Qy	182	PHYDLSMWTLIIQOTPCANGFVSLQAEVCGAFTDLPYRPDAVLVFCGAIAITLVGGQVKAP	241
Db	183	PHYDLSIIITTIQOTPCANGFVSLQVDVDGRVVDLPFVPGCVWVFCGAVAPLVSGGRIKAP	242
Qy	242	RHHVAAPREDQIAGSSRRTSVFFLEPRNADTFSPVPLARECGFVSLDGETATTFQDMIGGN	301
Db	243	RHQVASPSMQQRVGSRRTSVFFLEPRKPDFRFPVPLARASGLDVFDTGDTATTFGEWIGGN	302
Qy	302	YVNIR 306	
Db	303	YVNLR 307	

RESULT 11
US-09-791-537-94436

```

US-09-791-537-94436
: Sequence 9436, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STR
: TITLE OF INVENTION: METHODS OF USE THERE
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 94436
: LENGTH: 332
: TYPE: PRT
: ORGANISM: Acremonium chrysogenum
US-09-791-537-94436

```

Query Match	56.6%	Score 926.5	DB 22	Length 332
Best Local Similarity	56.7%	Pred. No. 7.1e-97		
Matches 177: Conservative	43	Mismatches 91	Indels 1	Gaps 1

Qy	1	MDTVPFTSLAELQOGLHQDFRRCRLDKGLFYLTDCGLTDTLTELKSAKDVLIDIFFEHGSE	60
Db	1	MTSKVPVFLDDLKSGKVLTLAEAVTTTKGIFYLTESGLVDDDDHTSARETCVDFFKNGSE	60
Qy	61	AEKRAVTSVPVPMRSGFTGLESESTAQITNTGYSYSDYSCMYSMGCTADNLPPSGDGFGRWT	120
Db	61	E EKRAVTLADRNRARFGSALEWESPAVVTETGKYSYSDYSCMIGIGNLFPNRGFEDVWQ	120
Qy	121	QYFDROYTASRAVAREVLRACTGEPDG-GVEAFDCBPLLRFRYPVOPVPHRSAAEOPLR	179
Db	121	DYFDNRYGAAKOVARAVLNSGAPLAGEDIDDFVECDPLLRFPVPEVDRVAAEEPLR	180
Qy	180	MAPHYDLSMTVLTIOQTPCANGFVSILQAEVGAFTDLPYRPDAVLVFCGAJATLVGTGQVK	239
Db	181	MGPHYDLSITILVHQTACANGFVSILQCEVDGEFVDLPTLPGAMVVFCGAVGTLATGSKVK	240
Qy	240	APRHVVAAPRRDQIAGSSRTSSVFLPRNADFTTSVPLARECGPDVLDGETATFQDWIG	299
Db	241	APKRVKSPGRDQVSSSRTSSVFLRPKPDFSFNVQQRSEGWGFNVIPISERTTFRWLG	300
Qy	300	GNTVNIIRRTSKA 311	
Db	301	GNTVNMRRDKPA 312	

```

RESULT 13
US-09-791-537-89857
; Sequence 89857, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89857
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces jumonjinensis
US-09-791-537-89857

Query Match          10.9%;   Score 179;   DB 22;   Length 329;
Best Local Similarity 24.9%;   Pred. No. 9.7e-11;

```


; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12239
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12239

Query Match 10.3%; Score 168; DB 29; Length 366;
Best Local Similarity 24.7%; Pred. No. 2.1e-09;
Matches 85; Conservative 41; Mismatches 158; Indels 60; Gaps 14;
QY 1 MDTTTFSLAELOQGHQ-----DFRCLRDKGLFYLTDCLTDLKSAKDLVIDRF 55
DB 28 MPRIVPVLDSRLSQASERRTFADLRASARDIGFFYLAGHGISWAEISEVLTASRQFF 87
QY 56 EHGSEAEKRAVTSVPVPMRGFTGLESESTAQITNTGYSYSDYSCYSMGTADNLFPSPGDF 115
DB 88 AL-PEADKLAIEVMKSSQFRGYTRAGELTK-----GREDWREQLDIGVERQAIAGPG 140
QY 116 GRIWTO-YFDRQYTAS-----RAVAREVLRATCTEPDGGVEAFDCEPLL 159
DB 141 TPATRLQGNQWPAALPDLKPALLAWQSKVTAVAIRLLKAFQAQSLDQPEDAF---DPI- 196
QY 160 RFRYFQVPEHR-SAEQPLR-----WAPHYDLSWVTLIOOTPCANGFVSLQAEVGG 210
DB 197 ----YSSEPNHRMKIVRYPGRDTGGDQGVGAHKDGGFLTLLO-----DDNKGQVYDYG 248
QY 211 AFTDLPVRPDVLFVCGAIAITLVGQGVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
DB 249 SWYDVDPFGTLVNVNIGELLELASNGYLRTATVHRVTPP-----AGVERISVPEFFSARLD 304
QY 271 FT-----FSVPLARECGFDVSLDGETATFODWIGGNVNIIRTS 309
DB 305 ATIPLLGLSEELAAQARGPAS-DPDNPLFRD-VGTNVLKSRLS 346

Search completed: February 3, 2004, 17:41:40
Job time : 176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 17:37:44 ; Search time 22 Seconds
(without alignments)
1088.984 Million cell updates/sec

Title: US-09-582-486-1

Perfect score: 1637

Sequence: 1 MDTVTFTSLAEQLQGLHQD.....ATFQDWIGGNYNVRTSKA 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374145 seqs, 77034267 residues

Total number of hits satisfying chosen parameters: 374145

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	174.5	10.7	345	6	US-10-679-063-16188
2	169	10.3	412	6	US-10-679-063-16144
3	160	9.8	342	6	US-10-679-063-22244
4	155	9.5	375	6	US-10-425-114A-58493
5	154	9.4	352	6	US-10-425-114A-63175
6	153	9.3	371	6	US-10-425-114A-62424
7	153	9.3	372	6	US-10-425-114A-53454
8	153	9.3	372	6	US-10-425-114A-72059
9	152	9.3	379	6	US-10-679-063-13321
10	148.5	9.1	389	6	US-10-481-381-4
11	137.5	8.4	335	6	US-10-679-063-9242
12	137.5	8.4	380	1	PCT-US03-24364-36
13	136.5	8.3	380	6	US-10-257-494A-16
14	135.5	8.3	405	6	US-10-739-930-7737
15	133	8.1	385	6	US-10-425-114A-66945
16	131.5	8.0	371	7	US-60-478-196-3232
17	131	8.0	356	6	US-10-425-114A-64811
18	131	8.0	356	6	US-10-739-930-7228
19	131	8.0	378	6	US-10-739-930-6723
20	131	8.0	378	6	US-10-257-494A-15
21	131	8.0	542	5	US-09-614-150A-11103
22	130.5	8.0	353	6	US-10-425-114A-61294
23	130.5	8.0	382	6	US-10-425-114A-66617
24	128.5	7.8	342	6	US-10-679-063-8477
25	128	7.8	346	6	US-10-679-063-15920
26	127	7.8	334	6	US-10-679-063-9241

ALIGNMENTS

RESULT 1

US-10-679-063-16188
; Sequence 16188, Application US/10679063
; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 16188
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Rhodospirillum rubrum
US-10-679-063-16188

Query Match 10.7%; Score 174.5; DB 6; Length 345;
Best Local Similarity 26.8%; Pred. No. 8e-09;
Matches 81; Conservative 32; Mismatches 134; Indels 55; Gaps 12;

Qy	6	PTFSL-AELQOGLHODEF-----RRLDKGLVLTDCGLTDTLTKSAKOLVIDPFEHSE	60
Db	13	PVLSLKAGEGEAARRDEFLATLRQAARDPCAFYLEGHGIDPSVIDQVEALSRRFPAL-PE	71
Qy	61	AEKRAVTSVPVTMRRTGFTGLESESTAQITNTGSYSYSCYSGMTADNLPSPGDFGRWT	120
Db	72	DEKRAIDMVNSPHFHGYTRVGAELTR-----GAPDREQLDIGSERPLLPGQPDTPAWA	125
Qy	121	QY-----FDRQVTA-SRAVAREVLRTGTEPDGGEAFIDC-----	155
Db	126	RLQGNQWPAALPDLRAAQLRLQAEALTAVALLERIALALGERAD----FFADLYEGGP	181
Qy	156	EPLLRFERYEPQVPEHRSASEQPLRMAPHYDLSMVTLLQOTPCANGFVSLQAEVGGFTDL	215
Db	182	DQLKTIIRYP-----GGAAGEGDOGVGPHKDSGLLTFLVLRDRG-----GLQIEREGHWDV	233
Qy	216	PVRPDVAVFCGAIATLVGTGGQVKAPRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSV	275
Db	234	PRPGTFVANNIGELLELATNGYLKATVHRVVSPPAD-----SDRLSTAFPL--CARLGSKV	287
Qy	276	PL 277	
Db	288	PL 289	

RESULT 2

US-10-679-063-16144

[illegible]

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT FILING DATE: 2003-04-28
CURRENT APPLICATION NUMBER: US/10/425,114A
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 63175
LENGTH: 352
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3062-030-F8_FLI.pep
US-10-425-114A-63175

Query Match 9.4%; Score 154; DB 6; Length 352;
Best Local Similarity 23.4%; Pred. No. 7.7e-07;
Matches 78; Conservative 49; Mismatches 119; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 41 TSATVSLPIVDLSLGRDEVRRRAILEAGKEIGFQVNVHGVSLAMQDMETVCOEPRF--L 98
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSMCYSMGTADNLPP 111
DB 99 PAEDKA-----GLYSEDTRATRIYSSMTFDTGCKEYWRDCLRLACS---PP 142
QY 112 S-GDGRIW-----TQYFDROYTASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 143 AVGDSAAAMPDKPRRLREVERFTVQTRGLGMEILRLCGLGLRDPYLEGDISGG--- 198
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEYGGAF 212
DB 199 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 238
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263
DB 239 LEVAYRGDWIRVPEVPGAFVNVFGCQLEVTNGILKSIHRVTN-----LGVARTTVAT 293
QY 264 FLRPNADFTFSVPLARECGFDVSLDGETATFQDW 297
DB 294 FIMPTTDCILIG-PAABFLSDNDPPCYRTLTFGDF 326

RESULT 6
US-10-425-114A-62424
Sequence 62424, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT FILING DATE: 2003-04-28
CURRENT APPLICATION NUMBER: US/10/425,114A
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 62424
LENGTH: 371
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3357-058-F2_FLI.pep
US-10-425-114A-62424

Query Match 9.3%; Score 153; DB 6; Length 371;

Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 60 TSATVSLPIVDLSLGRDEVRRRAILEAGKEIGFQVNVHGVSLAMQDMETVCOEPRF--L 117
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSMCYSMGTADNLPP 111
DB 118 PAEDKA-----GLYSEDTRATRIYSSMTFDTGCKEYWRDCLRLACS---PP 161
QY 112 S-GDGRIW-----TQYFDROYTASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 162 AVGDSAAAMPDKPRRLREVERFTVQTRGLGMEILRLCGLGLRDPYLEGDISGG--- 217
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEYGGAF 212
DB 218 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 257
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263
DB 258 LEVAYRGDWIRVPEVPGAFVNVFGCQLEVTNGILKSIHRVTN-----LGVARTTVAT 312
QY 264 FLRPNADFTFSVPLARECGFDVSLDGETATFQDW 297
DB 313 FIMPTTDCILIG-PAABFLSDNDPPCYRTLTFGDF 345

RESULT 7

US-10-425-114A-53454
Sequence 53454, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 53454
LENGTH: 372
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700169504_FLI.pep
US-10-425-114A-53454

Query Match 9.3%; Score 153; DB 6; Length 372;
Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
QY 4 TVPTFSLAELOQGLHODEFRCL-----RDKGLFYLTDCGLTDTTELKSAKDLVIDDFEHGS 59
DB 61 TSATVSLPIVDLSLGRDEVRRRAILEAGKEIGFQVNVHGVSLAMQDMETVCOEPRF--L 118
QY 60 EAEKRAVTSVPVPTMRGFTGLESESTAQIT-----NTGSYSDYSMCYSMGTADNLPP 111
DB 119 PAEDKA-----GLYSEDTRATRIYSSMTFDTGCKEYWRDCLRLACS---PP 162
QY 112 S-GDGRIW-----TQYFDROYTASRAVAREVLR-----ATGTEPD-----GGVEAF 152
DB 163 AVGDSAAAMPDKPRRLREVERFTVQTRGLGMEILRLCGLGLRDPYLEGDISGG--- 218
QY 153 LDCEPLLRFRYPQVPEHRSABEQPLRMAPHYDLSMTLIQOQPCANGFVSLQAEYGGAF 212
DB 219 ---DVLVHNHYPPCPDPNAT---LGLPPHCDRNLLTL-----LPSMWPG-- 258
QY 213 TDLPYR-----PDAVLVFCGAIATLVGGQVKAPRRHVAAPRRDQIAGSSRTSSVF 263

```

; SEQ ID NO 13321
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Fagus sylvatica
US-10-679-063-13321

Query Match
  9.3%; Score 152; DB 6; Length 379;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;

QY 30 GLFYLTDCLTDTTELKSAKDLVIDFFHSGEAEKRAVTSVPVPMRR-----GFTG- 79
Db 92 GFFLVNHGVDDKLIHAHAHQYIDYFFELPMSAQRA-----QRKVGSHCGYASSFTGR 144
QY 80 -----LESESTAQITNTGSDYSYCMVSMGTADNLFPSGDFGRITWTOYFDRQYTA 129
Db 145 FSKFLPWKETLSFRSSAQPDSSNIQDY-LCNTMG--EDFKP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGEVAFDCEPLLRYPFP--QVPEHRSAGEQPLRMAPHYDLS 187
Db 199 SLGIMELLGMSLVGSQGHYREFEENESIMRLNYPPCKP-----DLTLGTGPHCDPT 252
QY 188 MVTLIQOTPCANGFVSLQAEVGG--AFTDLPYR-----DAVLVFCGAIATLVGGQVKA 240
Db 253 SLTILHQD-----QVGGLOQVFDDEWRSITPNFNAFVNIGDTFMALSNGRYKS 301
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVP 276
Db 302 CLHRAVWNSK-----TPRKSALFFLCPKNDKVVSPP 332

RESULT 10
US-10-481-381-4
; Sequence 4, Application US/10481381
; GENERAL INFORMATION:
; APPLICANT: Okawa, Miho
; APPLICANT: Matsuo, Makoto
; APPLICANT: Ashikari, Motoyuki
; TITLE OF INVENTION: sd1 gene involved in plant semidwarfing and uses thereof
; FILE REFERENCE: SHZ-018US
; CURRENT APPLICATION NUMBER: US/10/481,381
; PRIOR FILING DATE: 2003-12-18
; PRIOR APPLICATION NUMBER: JP 2001-185128
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-481-381-4

Query Match
  9.1%; Score 148.5; DB 6; Length 389;
Best Local Similarity 24.8%; Pred. No. 2.9e-06;
Matches 64; Conservative 40; Mismatches 113; Indels 41; Gaps 10;

QY 30 GLFYLTDCLTDTTELKSAKDLVIDFFHSGEAEKRAVTSVPVPMRRGFTGLESE---STA 86
Db 94 GFFQVSEHGVDAALARAALDGSDFRLPLAEKRR--RRVFTVSGYTSAHADRFASKL 151
QY 87 QITNTGSDYSYCMVSMGTAD---NLPPSGDF---GRIWTOYFDRQYTSRAVAREVLR 139
Db 152 PKWETLSFGFHRAAAPVADVFSSTLGP--DFAPMGVYQKYCEMKELSLTIMELLEL 209
QY 140 ATGTEPDGGEVAFDCEPLLRYPFPQVPEHRSAGEQPLRMAPHYDLSMVTLIQOTPCAN 199
Db 210 SLGVEGYVREFFADSSSIMRCNYYPPCP-----PERTLGTGPHCDPTALTIL----- 258
QY 200 GFVSLQAEVGG--AFTDLPYR-----PDVAVFCGAIATLVGGQVKAAPRHVAAPRDQ 252
Db 259 ----LQDDVGGLEVLVDGEWRVSPVPGAMVINIGDTFMALSNGRYKSLHRAVWNSK 314
QY 253 IAGSSRTSSVFFLRPNAD 270

; SEQ ID NO 13321
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Fagus sylvatica
US-10-679-063-13321

Query Match
  9.3%; Score 153; DB 6; Length 372;
Best Local Similarity 23.4%; Pred. No. 1e-06;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;

QY 4 TVPTFSLAELOQGLHODEFRCL---RDKGLFYLTDCLTDTTELKSAKDLVIDFFHSGS 59
Db 61 TSATVSLPIVDLSGRDEVRRALEAGKEIGFFQVNVHGVSLQAMQDMETVCOEFFR--L 118
QY 60 EAEKRAVTSVPVPMRRGFTGLESESTAQIT-----NTGSDYSYCMVSMGTADNLFPP 111
Db 119 PAEDKA-----GLYSEDIGRATRIYSSMTFTGGEKYWRDCLRLACS---FP 162
QY 112 S-GDGRIM-----TOYFDROYTSRAVAREVLR-----ATGTEPD-----GGVEAF 152
Db 163 AVGDSAAANPKPRRLREVVERFTVQTRGLGMEILRLCEGLGLRDPYLEGDISGG--- 218
QY 153 LDCEPLLRYPFPQVPEHRSAGEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAEVGGAF 212
Db 219 ---DVLVNHVHPPCPDPNAT-----LGLPPHCDRNLLTL-----LPSWVPG-- 258
QY 213 TDLPYR-----PDVAVFCGAIATLVGGQVKAAPRHVAAPRRDQTAGSSRTSSVF 263
Db 259 LEVAYRGDMIRVEPVGAFVNVFGCQLEVVNTGILKSIEHRVTN-----LGVARITVAT 313
QY 264 FLRPNADFTFVPLARECGDFVSLDGETATFQDW 297
Db 314 FIMPTDCLIG-PAABFLSDNDNPPCYRTLTFGDF 346

RESULT 9
US-10-679-063-13321
; Sequence 13321, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
```

```

; SEQ ID NO 13321
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Fagus sylvatica
US-10-679-063-13321

Query Match
  9.3%; Score 152; DB 6; Length 379;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 67; Conservative 40; Mismatches 105; Indels 64; Gaps 12;

QY 30 GLFYLTDCLTDTTELKSAKDLVIDFFHSGEAEKRAVTSVPVPMRR-----GFTG- 79
Db 92 GFFLVNHGVDDKLIHAHAHQYIDYFFELPMSAQRA-----QRKVGSHCGYASSFTGR 144
QY 80 -----LESESTAQITNTGSDYSYCMVSMGTADNLFPSGDFGRITWTOYFDRQYTA 129
Db 145 FSKFLPWKETLSFRSSAQPDSSNIQDY-LCNTMG--EDFKP---FGKYQDYCDAMSTL 198
QY 130 SRVAREVLRTATGTEPDGGEVAFDCEPLLRYPFP--QVPEHRSAGEQPLRMAPHYDLS 187
Db 199 SLGIMELLGMSLVGSQGHYREFEENESIMRLNYPPCKP-----DLTLGTGPHCDPT 252
QY 188 MVTLIQOTPCANGFVSLQAEVGG--AFTDLPYR-----DAVLVFCGAIATLVGGQVKA 240
Db 253 SLTILHQD-----QVGGLOQVFDDEWRSITPNFNAFVNIGDTFMALSNGRYKS 301
QY 241 PRHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVP 276
Db 302 CLHRAVWNSK-----TPRKSALFFLCPKNDKVVSPP 332

RESULT 10
US-10-481-381-4
; Sequence 4, Application US/10481381
; GENERAL INFORMATION:
; APPLICANT: Okawa, Miho
; APPLICANT: Matsuo, Makoto
; APPLICANT: Ashikari, Motoyuki
; TITLE OF INVENTION: sd1 gene involved in plant semidwarfing and uses thereof
; FILE REFERENCE: SHZ-018US
; CURRENT APPLICATION NUMBER: US/10/481,381
; PRIOR FILING DATE: 2003-12-18
; PRIOR APPLICATION NUMBER: JP 2001-185128
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-481-381-4

Query Match
  9.1%; Score 148.5; DB 6; Length 389;
Best Local Similarity 24.8%; Pred. No. 2.9e-06;
Matches 64; Conservative 40; Mismatches 113; Indels 41; Gaps 10;

QY 30 GLFYLTDCLTDTTELKSAKDLVIDFFHSGEAEKRAVTSVPVPMRRGFTGLESE---STA 86
Db 94 GFFQVSEHGVDAALARAALDGSDFRLPLAEKRR--RRVFTVSGYTSAHADRFASKL 151
QY 87 QITNTGSDYSYCMVSMGTAD---NLPPSGDF---GRIWTOYFDRQYTSRAVAREVLR 139
Db 152 PKWETLSFGFHRAAAPVADVFSSTLGP--DFAPMGVYQKYCEMKELSLTIMELLEL 209
QY 140 ATGTEPDGGEVAFDCEPLLRYPFPQVPEHRSAGEQPLRMAPHYDLSMVTLIQOTPCAN 199
Db 210 SLGVEGYVREFFADSSSIMRCNYYPPCP-----PERTLGTGPHCDPTALTIL----- 258
QY 200 GFVSLQAEVGG--AFTDLPYR-----PDVAVFCGAIATLVGGQVKAAPRHVAAPRDQ 252
Db 259 ----LQDDVGGLEVLVDGEWRVSPVPGAMVINIGDTFMALSNGRYKSLHRAVWNSK 314
QY 253 IAGSSRTSSVFFLRPNAD 270
```

```

Db      |   ||||||   |   |
       315 -----RRSLAFFLCPRD 327


RESULT 11
US-10-679-063-9242
; Sequence 9242, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 9242
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Citrus unshiu
US-10-679-063-9242


Query Match          8.4%; Score 137.5; DB 6; Length 335;
Best Local Similarity 21.8%; Pred.No.2.8e-05;
Matches            63; Conservative    47; Mismatches 124; Indels     55; Gaps    10;

Qy        5 VPFTS LAELQQ LQHDFRRC LRDKGL FYLTDC GLTDT ELKSAKD LVIDFEEHSGSEAEKR 64
Db         43 IPTTLDDP QVDDR LVRSIAEASREWGIFQVTNHGIPSDLIC KLOAVGKEFFPELPQE-EKE 101
Qy        65 AVTSFPVTMRRGFTGLESESTAQTITNGSYSDYSMGCTADNLNLFPSGDGFRIWT----- 120
Db         102 VYSPRADA--KVQG YGTKLQ KEVGKKSWVDHL-----FH RVWPSSI 143
Qy        121 --QVFDRQYTASRAVAREVLR-----ATGETPDGGV--BAF--LDCEPLLRF 161
Db         144 NYREFWPKNPFSYRAVNIEAYAKMYREVVDKLGFTYL SLGLGVSEGVLK EAAAGGDDI EYMLKI 203
Qy        162 RYPQVPVEHRSAEQRLMAPHYDLMSMTLLQO TPCANGFVS LOA EVCGAGFT DLPPRPDA 221
Db         204 NYPEPCR----PDALGVNAHTDSALTVI-----VPNVEPGLVQFKDDRWDIKAPIPA 255
Qy        222 VLVEFGAIATLV TGQVKAPRHVAA PRRDGIAGSSRTSSVFFLRNAD 270
Db         256 LVIHIGDOIETLS NGKYKA VLIRTTNVK-----DKTRMSWFVLEPPAD 299
```

```

RESULT 12
PCT-US03-24364-36
; Sequence 36, Application PC/TUS0324364
; GENERAL INFORMATION:
; APPLICANT: BASF PLANT SCIENCE GMBH
; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS IV
; FILE REFERENCE: 16313-0236
; CURRENT APPLICATION NUMBER: PCT/US03/24364
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US 60/400,803
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
PCT-US03-24364-36

```

```

Qy 59 SEAEKRAVTSVPVTMR-----GFTG-LESESTAQITNTGSDYSYCMYSGMTADN 108
Db 117 PACEKQ-----KAQRKMGSSGYASVFGRFSSKLPWKETLSFKFSPKEIKHSQTVKD- 169
Qy 109 LFPS-----GDFGRITWQVDFRQYATSAVAREVLRACTEPDGGVEAFLDCEPLLRF 161
Db 170 -FVSKMCDGYEDFGKYQVEAEAMNTLSLKIMELLGMSLGVERRYPKEFFEDSDSI 228
Qy 162 RYFPQVPEHRSABEQPIRMAPHYDLSWTLIQOTPCANGFVSLQAEVGGATDILPYRPDA 221
Db 229 NYTPQCKQ----PBLALGTGPHCDFTSLTILHODQVG---GLQVFDNKMQSIPPNPHA 280
Qy 222 VLNVCGAIAITVTCGGQVAKPRHHVAAPRRDQIAGSSRTSSVFFILRPNADFTFSVP 276
Db 281 FVNIGDTFWALTNTRYKSLCHRAVNSRE-----RKTFAFFLCPKGEKVVKPP 330

RESULT 13
US-10-257-494A-16
; Sequence 16. Application US/10257494A
; GENERAL INFORMATION:
; APPLICANT: AN, GYN-HEUNG
; APPLICANT: KANG, HOUG-GYU
; APPLICANT: JUN, SUNG-HOON
; APPLICANT: KIM, JUNEUL
; APPLICANT: JUNG, KI-HWAN
; APPLICANT: PARK, YONG-JOO
; APPLICANT: LEE, SANG-YEB
; TITLE OF INVENTION: A PROMOTER CY20oxP WHICH REGULATES THE INTEGRUMENT-SPECIFIC
; TITLE OF INVENTION: EXPRESSION OF A GIBBERELLIN 20-OXIDASE GENE IN DEVELOPING
; TITLE OF INVENTION: SEEDS OF WATERMELON AND A METHOD FOR GENERATING SEEDLESS
; TITLE OF INVENTION: FRUITS USING THE PROMOTER
; FILE REFERENCE: 7022-0003
; CURRENT APPLICATION NUMBER: US/10/257,494A
; CURRENT FILING DATE: 2003-08-16
; PRIOR APPLICATION NUMBER: PCT/KR00/01127
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: KR 2000/18483
; PRIOR FILING DATE: 2000-04-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-257-494A-16

```

```

Query Match      8.3%; Score 136.5; DB 6; Length 380;
Best Local Similarity 23.4%; Pred. No. 4.1e-05;
Matches 69; Conservative 38; Mismatches 143; Indels 45; Gaps 11;

Qy  5  VPTSLAELOQG---LHQDSFR---RCIRDKGLFYLTDCGLTDTLXSAKDVLVDFPHG  58
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  58  VPLDLAGFLSGDCLASEATRLVSKAATKHGFPLITHGV-DESLSLRAYLHMDSPFKA  116
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy  59  SEAEKRAVTPVPTMR-----GFTG-LESESTAQIITNGSYSDSYSCYSGMTADN  108
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  117  PACEKQ-----KAQRKMGSSGVASSFVGRFSSKLPKWETLSFKFSPEEKIHQTVKD-  169
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy  109  LFPS-----GDFGRITWQYDFRDQYTA$RAVAREVL$RATGTEPDGGV$EAF$DCEPLLR  161
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  170  -FVSKMGDGYEDGKYVQYEA$AMNTLSLKIMELLGMSLGVERRYKPEFFEDSDSIFRL  228
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy  162  RYFQVPEHRSABEQPLRMAPHYDLSVMTLIQOFTPCANGFVSLQAEVGGATDLPYR$DA  221
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  229  NYPQCKQ---PELALGTGPHCDPTSLITLHQDQVG---GLQVFDNKNQCSIPPNPHA  280
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy  222  VLVFCGAITLVTCGVKAPHHVAA$RRQJAGSSRTSSVFELRPNADFTFSVP  276
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  281  FV$NIGDTFMALTNGRYK$CLHRAV$N$RE-----RKTFAFFLCPKGEKVVKPP  330
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 14
US-10-739-930-7737
; Sequence 7737, Application US/10739930
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7737
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(405)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C24992_1.p
US-10-739-930-7737

Query Match 8.3%; Score 135.5; DB 6; Length 405;
Best Local Similarity 24.2%; Pred. No. 5.5e-05;
Matches 61; Conservative 41; Mismatches 123; Indels 27; Gaps 8;
QY 30 GLFYLDCGLTDLTELAKDLVDFPHGSEAEKRAVTSVPVPMRGFTGLESESTA--- 86
Db 113 GPFQVCGHGVDAALGRAALDAGSDFFRLPLAEKORA--RRVPGTVSGYTSAHADRFPAKL 170
QY 87 --OITWTGYSYDYS-----MCYSMTADNLFPSGDFGRITWQYDFDRQYTSARAVAREVLR 139
Db 171 PKKETLSFGHGDGAXPVVDVYXVGLGQDFE--PMGWYQRYCEEMKELSLTIMELLEL 228
QY 140 ATCTEPDGGV-EAFLDCEPLLRFPQVPEHRSABEQPLRMAPHYDLSMVTLIQOTPCA 198
Db 229 SLGVELRGYVREFEFSRSMRCNYPPCPPE---PERTLGTGPHCDPTALTILLQDXDV 284
QY 199 NGFVSLOAEVGGAFDLPYRPDAVLVFCGAIATLVGGQVKAPRHHVAAPRRDQIAGSSR 258
Db 285 GG---LEVLDGEMRPVPPVPGAMVINIGDTFMAISNGRYKSLHRAVVNQ-----RAR 336
QY 259 TSSVFFLRPNAD 270
Db 337 RSLAFFLCRPED 348

RESULT 15
US-10-425-114A-66945
; Sequence 66945, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 66945
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4756-039-A6_FLI.pgp
US-10-425-114A-66945

Query Match 8.1%; Score 133; DB 6; Length 385;

Best Local Similarity 21.9%; Pred. No. 9e-05;
Matches 61; Conservative 43; Mismatches 140; Indels 34; Gaps 9;
QY 5 VPTFSLAELQQ---GLHODEFRERCLRDKGLFYLTCGLTDTLTELKSAKDILVIDFFEHGSEA 61
Db 67 IPVIDVTTELQHDHVDGL--DKLRLACEEWGFFOVVNHGIAHELLEDEMERLAREFFMLPLEE 124
QY 62 EKRAVTSVPVPMRGFTGLESESTAQITNTGYSYDYSYCMYSMT-----ADNLFPSS--G 113
Db 125 KEKYPMAP-----GGIQGYGHAFVFSQDKLWCNNMALGVPEPASIRQPLNPTAPA 176
QY 114 DFGRIWTOYDFDRQYTSARAVAREVLRATGTEPDGGVEAFDCEPLLRFPYFQVPEHRS 173
Db 177 GFGETLEAYSAEVSELCCRLLARIAETLGLAPATTFADMFGAEAVQVRMNFYPPCPR---- 232
QY 174 BEQPLRMAPHYDLSMVTLIQOTPCANGFVSLQAE--VGGAFDLPYRPDAVLVFCGAIATL 232
Db 233 PLVVMGLSAHSDGSVAVTLQQDV---GCAGLQVRGKGAWVPVHPVPHALVVNIGUTLEV 289
QY 233 VTGGQVKAPRHHVAAPRRDQIAGSSRTSSVFFLRPNAD 270
Db 290 LTNGRYKSVEHRAVAN-----GEQDLSLVVTFYAPAYD 322

Search completed: February 3, 2004, 17:43:36
Job time : 24 secs